

# *Hiding Message to BMP Image With LSB Insertion Using Python*

Anwar Ramadha

School of Electrical Engineering and Informatics  
Bandung Institute of Technology  
Jl. Ganecha 10 Bandung, Indonesia 40132  
Email : anwar.ramadha@gmail.com

**Abstract**—There are so many programming language in the world. From the language that need to be compiled and not like python do. Beside it, today, the secrecy of data become very important. So, many experts try to find method to hide message. As we know that are encryption and the second one maybe not familiar, steganography. In this paper, writer will show how to hide message to a BMP image file with python.

**Keywords**—Python, Steganography, BMP image

## I. INTRODUCTION

Since the invention of computer by Charles Babbage in 1882, computer needs an instruction to perform a task. So, a method called programming language was created. The computer language has two stage, the first major language and the second major language.

Over sixty years after that, In 1949 Von Neumann create the first programming language based on 0 and 1. Then, in 1951, Grace Hopper wrote the first compiler. Compiler is the program to convert programming language to 0 and 1 so that the computer can understand it.

As technology develops, the programming language more and more springing up. From the language that needs compiler and the second one that don't need to be compiled. Just run it. It called interpreted programming language.

When a machine execute instruction of a program, it will use machine language which is composed from 0 and 1. So the compiled language and interpreted language has same method. They will change their syntax to machine language.

Using the interpreted language, we can get advantages and disadvantages at the some time and vice versa. The advantages of interpreted language are easier to implement, no need compilation stage, and can be more convenient for dynamic language. But the disadvantage is the interpreted language will be more slower because the interpreter will change the instruction to machine language before run. It will take more time than compiled program, because it already in machine language. The example of popular interpreted language are PHP and Python. The second one will be explained.

This paper will use a technique to hide message to an image which use LSB (Least Significant Bit) to store message to explain some python's syntax.

There are many algorithm to hide message to an image. But in this paper, writer just use naive technique to implement this and writer would not give detail for it. This paper just show little use of python.

## II. PYTHON



Figure 1 Python's logo

Guido Van Rossum published first version of python language (version 0.9.0) in February 1991. This release included error handling, exception, function, core data type like integer, str and many more. Python has object oriented at the first release.

After that, python 1.0 released in January 1994 and bring new feature like lambda, map, filter, and reduce.

Six and half years after python 1.0 released, in October 2000 python 2.0 introduced to public. In this version python get improvement and some feature like full garbage collector and support unicode has been added.

For another 8 years, python released version 2.x before python 3.0 introduced. But python 3.0 is not backward compatible with version 2.x. The emphasis in Python 3 had been on the removal of duplicate programming constructs and modules.

Why this language named python? Like the name of one snake species, right? Many people was wrong about it. Some people said that the logo of python language is two snakes. The origin name come from british humour. The creator of python, Rossum, G. V. (1996) wrote "Over six years ago, in December 1989, I was looking for a 'hobby' programming project that

would keep me occupied during the week around Christmas. My office ... would be closed, but I had a home computer, and not much else on my hands. I decided to write an interpreter for the new scripting language I had been thinking about lately: a descendant of ABC that would appeal to Unix/C hackers. I chose Python as a working title for the project, being in a slightly irreverent mood (and a big fan of Monty Python's Flying Circus)."

When python compared with other programming language like java, it have advantages and disadvantages. The java language is faster than python when executed. But it has 3-4 times longer in the script. This differences can be attributed to python's built-in high-level data type and its dynamic typing. The python's programmer would not waste their time for declaring type of variable. But it is the disadvantage too. Because the machine will work harder when run time to define the type of variable. Not same in java program which was variable has declared and the compiler was define it. So the java program can run efficiently. If python compared with C++, it will has 5-10 shorter than C++. Anecdotal evidence suggest that one python programmer can finish a task in two month what two C++ programmer can't complete in a year.

### III. STEGANOGRAPHY

There are large number of steganography method that most of us are familiar with. With computer and network, there are many way to hide message, such as convert channel, hidden text within web page, etc.

Steganography today, allow us to hide large amount of message within image and audio file. These form often conjunction with cryptography so that the message doubly protected. First it is encrypted and hiding it to a media like image or audio and finding the message in this media then decrypt it. Yeah, the word that say a picture more than a thousand word literally true for computer scientist.

There is very basic formula of steganography, describing the pieces of steganography process :

cover\_medium + hidden\_message + stego\_key (that is optional) = stego\_medium.

Cover medium is the file which we will hide a message or the hidden message. Than if we want so that the message more secure, we can encrypt the hidden message before insert it to cover medium. It called stego key. But in this paper, writer just insert the message directly to an image.

Before discuss how information hide to an image, writer will give short explanation how the image is stored.

Image file contains binary representation of the color or light intensity of each pixel. Images typically 8-bit until 24-bit color. When using 8-bit color, the image just have number 0 - 254 to represent RGB color. For a 24-bit color image has much better set of color. In Hypertext Markup Language (HTML) often use six hexadecimal digits. Each pairs of digit represent RGB in an image. For example, the orange color will represent as red 100% (decimal 255, hex FF), 50% green (decimal 127, hex 7F), and no blue (decimal 0, hex 0). This form will become

0xFF7F00 in complete six digits hexadecimal. The BMP file image is 8-bit color depth.

The size of an image depend on the number of pixels and the granularity of the color definition. An image that has 640 x 480 pixel in 8-bit color typically has 307 KB (640 x 480 pixels). On the other hand, the image that has 640 x 480 pixels and 24-bit color have 880 KB (640 x 480 x 3 bytes).

The simplest way to hide message within an image is using least significant bit (LSB) insertion. In this method, we will take the binary representation of hidden message and insert it for each bytes to cover image.

For example we have three adjacent pixels (8-bytes) with the following RGB encoding :

11000101	01101001	10110110
10111000	00110101	01110001
10011101	11100010	

Figure 2 Image data before LSB insertion

Now, we will insert a message, for example we want to insert 's' letter that has 01110011 in binary. For each number in binary representation of 's' letter will insert to LSB of image data. The following illustration will make more clear understanding.

110001 <b>00</b>	0110100 <b>1</b>	101101 <b>11</b>
1011100 <b>1</b>	0011010 <b>0</b>	0111000 <b>0</b>
1001110 <b>1</b>	111000 <b>11</b>	

Figure 3 Image data after LSB insertion

We can see above that the LSB of the pieces pixels of an image can change or not after modification.

### IV. IMPLEMENTATION

The first thing that we have learn is to read the documentation of python. Because python has differences from other popular language like java and the same type of interpreted language, PHP. But it is really simple language which writer use. But we have to strict about the indentation of code. If not, you'll get error message.

Okay, let's see how to implement the LSB insertion to hide message in BMP image with python.

First, we have to read data from image file. Python has simple method to open all type of file. if we using java, it will be different. The data which obtained from file reading, will have decimal value. We have to change it to binary. But the problem is the result of this transformation is it has 'b' letter in

this result. Such as we have 010011101, the transformation will be 0b10011101. We must remove the 'b' letter.

After that we can get the input data from user, but in this project, writer just make simple program that will hide the message with string value and using std in not read from file. You can place this step before or after you read the image data. but it be good choice if you place first because your computer would not save the image data in buffer longer. The message variable has initialized with the Start of Text (STX) character with '\x02' value. it will be useful when we extract the message from image. If the user finish to insert message, add End of Text (ETX) character that has '\x03' value. The next step is convert it to binary character by character. It has same problem with image reading step, we must remove 'b' letter.

If cover image and the hidden message obtained, we will iterate for each bit in each bytes of hidden message and insert it to the LSB. You can see how it works at the steganography session. But you have to start insertion from index 60 in image data. Because the 0-59 data is the header of BMP file. If you modify it, you will get an error while create the image back.

After insertion complete, we must create new image from this data. Before it, we must change image binary data to decimal again. If the data was successfully changed, just write the data to the file, hope your project completed successfully.

To get the hidden message back, we must load the image and do like the first step. After that, find the STX in loaded image. If the STX character found means that the image contain hidden message and you must continue the iteration, if not, stop the iteration and show the error message. in this step, you must gather the last character in binary data of image file and convert it to a character.

Writer has made to version of this program, first is using python, and the second one is using java. Using python has 92 line shorter than java. You can see writer's code in writer repository at <http://github.com/anwarramadha/steganography>.

## V. RESULT

in this section, we will see the result of steganography implementation. We will use spongebob.bmp. Most article said that lena picture is the best choice because it has color distribution equally. But writer don't want to share the picture of woman without hijab. Okay, let's see the result.

Okay. let's see how to implement the LSB insertion to hide message in BMP image with python.



Figure 4 image without hidden message



Figure 5 Image with hidden message

If we see, the two picture above are not different at all. Because human eye can't see the changed color gradation in this pictures.

The second picture has a hidden message which can extract. The following picture will show you what is in this picture.

```
anwar@anwar ~/Desktop/sosif $ python LSB_Hide_Message.py
Select option
1. Create hidden message
2. Open hidden message
3. Exit
Your choice : 2
Open secret image (only bmp image can be used) : spongebob-stego.bmp

=====
Secret message :
secret
=====
```

Figure 6 Extracted message

## VI. CONCLUSION

Along with technology advances, the secrecy of data become very important. There are many way to protect data content such as cryptography and steganography.

Choosing the language to develop it is very important too and using python as your choice can reduce time consuming to write script.

## REFERENCES

- [1] K our, J., Verma, D. (2014). Steganography : Hiding data within data. *Journal of Steganography*.3(5). Retrieved May 4, 2017, from [https://www.ermt.net/docs/papers/Volume\\_3/5\\_May2014/V3N5-190.pdf](https://www.ermt.net/docs/papers/Volume_3/5_May2014/V3N5-190.pdf)
- [2] Hystory of Python. (2003, January 13). Retrieved May 4, 2017, from [http://www.pythoncourse.eu/python3\\_history\\_and\\_philosophy.php](http://www.pythoncourse.eu/python3_history_and_philosophy.php)
- [3] Python documentation. (2017, April 30). Retrieved May 4, 2017, from <https://docs.python.org/3/>

STATEMENT

I hereby declare that the paper I am writing is my own, not an adaptation, or a translation from someone else's paper, and not plagiarism.

Bandung, May 4, 2017

A handwritten signature in black ink, consisting of a large, stylized letter 'A' followed by a vertical line and some smaller, less distinct characters.

Anwar Ramadha