IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 9, September 2016
ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

The Model and Method of Data Theft Based on Image Camouflage

Yu Tang

School of Electronic & Information Engineering, Jiangsu University of Science and Technology,
Zheng Jiang, Jiang Su, China

Abstract

Image camouflage technique is disguising the text data by stealing to an image file, in order to achieve file type camouflage and guarantee concealment of information theft. First, this paper introduces the structures of the theft model, and then introduces BMP image camouflage methods and JPEG image camouflage method. In order to improve the security of theft of data transfer, Finally camouflage image is uploaded to network disk. By capturing data packets uploaded, we analysis the experiment and the final results show that: the data theft method which is based on image camouflage has the concealed, confusing, and high security features and it can guarantee the delivery of data theft effectively.

Keywords: *BMP Image Camouflage, JPEG Image Camouflage, Network disk*

1. Introduction

With the widespread deployment of tools for Internet regulation and the increasing awareness of Internet security, more and more eavesdroppers communicate by camouflaging data snugly to effectively avoid discovery, interception, correlation or analysis of the Internet traffic. And the Internet traffic is more concealed, confusing, anti-intercepted and anti-analytical after being disguised [1].

Data camouflage [2] is also called data hiding, which means to deliver messages by hiding them in the public non-confidential document. So that, the observer and monitoring system is unable to detect the presence of some specific information. The specific information may include characters, password or images while public document means normal text files, digital images, audio or video, etc. Compared to encryption technique, data camouflage makes message invisible because the secret information is embedded into a digital vector before communication.

In this paper, the author camouflages text messages into different types of image files to ensure the diversity of files in the process of delivering messages. The communication flow is more concealed and effective after being camouflaged. In order to improve the security for data transmission, the popular cloud technique is adopted. The receiver can download the data from some specific cloud storage which ensures that data-thief won't communicate with the receiver directly so that stolen data will be more unnoticeable, secure and interference immune. In conclusion, the author establishes a data theft model based on image camouflage for data theft.

2. Model of Data Theft

The model discussed in this paper is shown in picture 1 as followed. After stealing the original data, the data thief needs to deliver them to someone else. First of all, the data thief camouflages original data behind images in the form of JPEG or BMP which makes the data invisible. And then upload the images into cloud storage. In order to get the stolen data, the data receiver must download all the files related from the cloud and extract the hidden original data from the images.

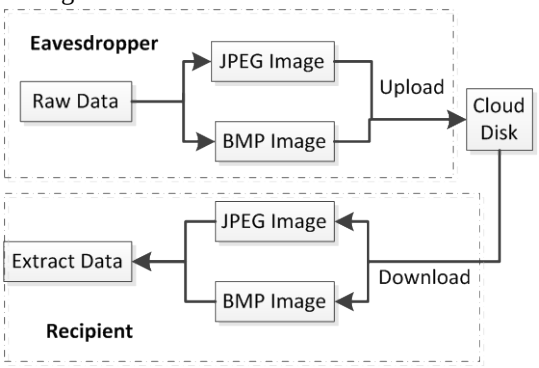

Fig. 1 Theft model based on the camouflage image

3. Method of Camouflage Image

Most of the conventional image camouflage[3] method can be seen as a replace system, these methods try to replace a redundant portion of the image information into the secret information, so that we achieve the purpose of encoding secret information. If the receiver knows the position of the secret information which is embedded, he can extract the secret information. These methods include using LSB encoding, image processing or image compression algorithms of feature (e.g., luminance) to be modified. The main drawback of this method is to modify the camouflage carrier has considerable vulnerability, while the carrier has little capacity to store information [4].

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 9, September 2016
ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

In this paper, we mainly use the JPEG and BMP formats as a carrier. We encode the theft information which will be transmitted as an image data block, and then draw an image according to the data in the block. In order to realize the theft of data disguised as an image, so that the data can not easily be found and this method has the main features of large storage capacity.

3.1 BMP Camouflage Image Method

BMP (Bitmap-File) image[5] files are also called bitmap files, bitmap represents that an image is divided into grids, each grid point called pixels, each pixel has its own RGB values, that is, a dot image is composed of a series of pixels. Bitmap file format supports four RLB and eight-bit and 24-bit code. This article is used in 24-bit format. Structure features of 24 BMP image[6] files are as follows: each file can only store one uncompressed color image file. The file consists of 54 bytes of data segments, which contains the bit file type, size, image size, and print format. Starting from the 55 bytes, the file is the image data of the file, the order data start from the lower left corner of the image, the order is from the left to right and from the button to up. Consecutive three bytes describe the color information of a pixel of the image, the three bytes represent the blue, green and red tricolor concentration in this pixel. The theft information needed to be hidden is stored in the data of the image.

Bitmap file consists of the file header, information header, image data, which is shown in Figure 2, with an example to explain:

```
          0  1  2  3  4  5  6  7  8  9  a  b  c  d  e  f
00000000h: 42 4D 9E 40 09 00 00 00 00 00 36 00 00 00 28 00 ;
00000010h: 00 00 F7 01 00 00 91 01 00 00 01 00 18 00 00 00 ;
00000020h: 00 00 68 40 09 00 C4 0E 00 00 C4 0E 00 00 00 00 ;
00000030h: 00 00 00 00 00 00 85 AC BB 84 B5 C5 7A BE D1 6F ;
00000040h: B8 CE 74 B3 CF 4C 85 A4 00 33 52 1A 5A 78 6C B9 ;
```

Fig. 2 BMP file format example

Part 1 to Part 4 (red line division) is a bitmap file header, '42 4D' is the label of the BMP files, '9E 40 09 00' is the size of the file in types.

Part 5 to part 15 (start part is divided by blue lines) is a bitmap information header about data and detailed information about BMP size. In the data, 'F7 01 00 00' is the width of BMP and '91 01 00 00'is the height which are all in pixels.

Part 16 (all data behind the green line division) is the image data, that is the storage of the theft data. When we generate BMP file, please note this "scan line" concept, the scan line refers to the size of the image scanning line of pixels in the memory bytes of data, which is also the size of the image scanning lines. The size depends on the number of colors of the image and the image width in pixels. Scanning the BMP format file[7] also has a very important provision requires that each scanning line must be divided by four. If the line image bytes can not be divided by four, we need to fill the zero to achieve the provisions.

Based on the above analysis, this paper converts the txt file of 15000 bytes to the binary stream. Binary data stream of the file is as an image data BMP files and generate it to BMP file. This BMP file is 100 pixels length and 50 pixels width. The entire BMP file size is 15054 bytes. This BMP file is amplified like Figure3, each pixel holds three bytes of data. Files of different sizes can lead to different length and width of the image size, achieving the camouflage from txt file to BMP file, camouflage method has the characteristics of large storage capacity.

Fig. 3 BMP image after camouflage

After BMP images was downloaded from the network disk, Recipient extract the file in the image data part which is the hidden theft information. Then, we convert the image data stored by a binary stream into the txt file and recover it to the original document.

3.2 JPEG Image Camouflage Method

JPEG[8] is the Joint Expert Group image abbreviation, the Expert Group developed algorithm called JPEG algorithm, and became the first international standard color, grayscale, still image internationally, so-called JPEG standard. LEG is currently developed two basic compression algorithms: one is DCT (Discrete Cosine Transform)[9] based on the lossy image compression algorithm; another is forecasting technology-based lossless compression algorithm. The lossy compression due to the compression process, remove the redundant information in the image, after compression and original images compared to the naked eye can not tell the difference between the two, we use the DCT compression algorithm.

JPEG file[10] consists of eight parts, each part of the tag is two bytes, the first byte is set to 0xFF, 0xFF allow multiple refilling in front, with a final standard. The main part of the file 8 contains: start of the file is marked as FF D8; APPO (image identification information) is labeled FF EO; APPn marker, where n = 1 ~ 15, corresponding to the value 0xE1 ~ 0xEF; one or more quantization table DQT., value 0xD8; frame image began SOFO, value of 0xCO; one or more Huffman table DHT, a value of 0Xc4; scanning start SOS, a value of 0xDA; compressed image data; image end EOI, the value is 0xD9. In the compressed data portion of the file is stored in the main drawing data, namely data storage theft.

According to the analysis of the above file formats, this article will be a size of 15,000 bytes txt file converted to the corresponding binary stream binary stream data of the file stored in the image data compression section JPEG data, set the image width and height, respectively, are 255 pixels, generating the JPEG image, shown in Figure 4, the entire JPEG file size is 15625 bytes.

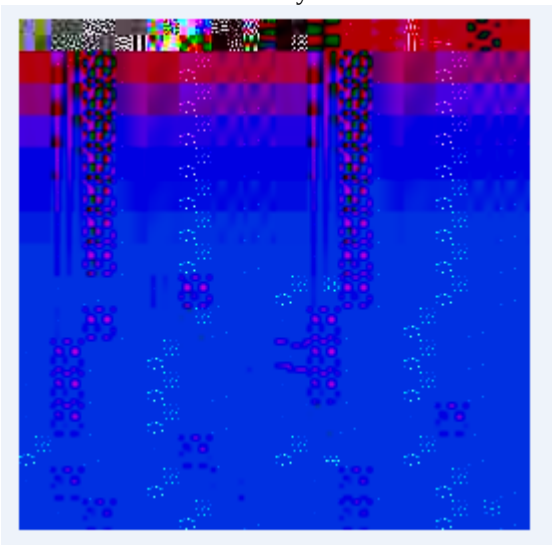

Fig 4 JPEG image after camouflage

The process of extracting the secret information is essentially the inverse process of embedding information. Since the theft information is stored in the compressed data of image when the information is embedded. When receiver downloads the JPEG images from a network disk, as long as extracting the binary data of the compression section which is of the image date, we then restore it to its original txt file.

4. Analysis of Results

In order to realize and demonstrate the theft model presented in this paper, the data theft interface like figure five which was based on camouflage image was designed to displaying the various functions of the model in this interface. First, we select a file stored locally, and then disguise them to any kind of image format and save the camouflage file in the local. Multiple theft files respectively repeat this procedure, and then we upload camouflage files to cloud disk which was agreed in advance.

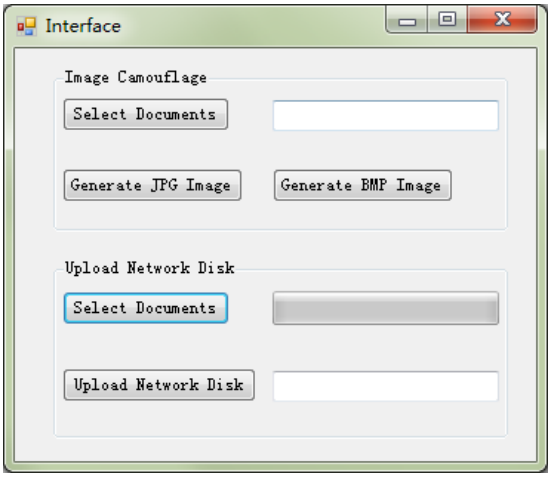

Fig. 5 Theft interface based on image camouflage

In order to verify the safety and concealment of this method, we performed several experiments. We upload BMP image like figure 3 and JPEG image like figure 4 to network disk, and then we use Wireshark tool to capture data packets for analyzing. We selected uploaded packets through the IP address of the machine. Figure 6 is a part of the data package for BMP. The size of the stream file is 15054 bytes which is the size of the Figure 3 BMP image file. 42 4d is the BMP header flag and the contents of the file are not visible, so the covert theft data has high hidden.

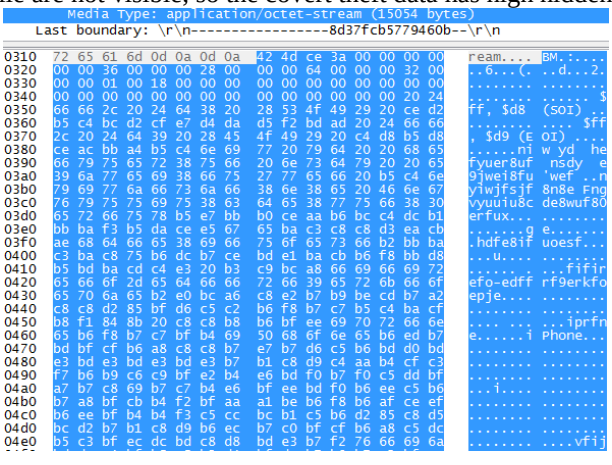

Fig. 6 BMP data packet uploaded

Figure 7 is a part of the data packet which is captured JPEG files uploaded. The file size is 15625 bytes stream which is the size of the JPEG image of figure 4. FF D8 is the header flag of the JPEG file. The file contents are also not visible, and ensure that the covert theft data has high hidden.

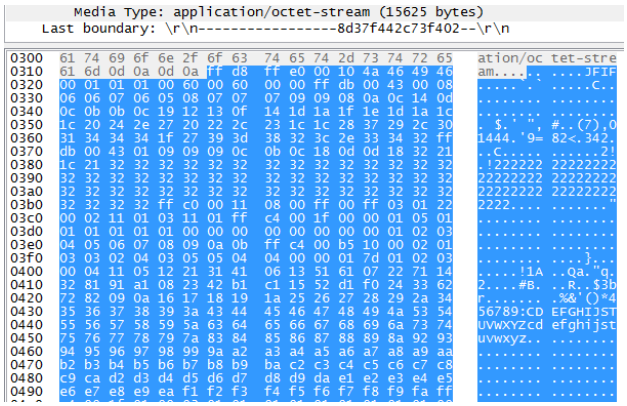

Fig. 7 JPEG data packet uploaded

To effectively avoid traffic discovery, capture and analysis, Eavesdropper disguises the data as BMP format files and JPEG files. Firstly, it achieves the conversion of network traffic, so that it effectively resists detect statistical traffic analysis. Secondly, file types after disguising have characteristics of diversity and highly Confusing. Meanwhile, Both sides eavesdropper are not directly communicating, but to upload the data to the network disk which is appointed in advance. With the network disk as a springboard to the middle, both the identity information has been hidden which ensures the concealment during the theft information transmission process and security.

5. Conclusion

The proposed data theft method which is based on image camouflage can effectively improve the security of the theft data. The detailed description of the BMP image camouflage methods and JPEG image camouflage shows that the theft of data disguised has the capacity, invisibility, safety and other important features. Meanwhile, the theft model combines today's popular cloud disk technology, uploading the disguised files to the network disk, thereby improving the safety of the information transmission process. The rapid development of network technology today, the method proposed in the file is an effective and feasible way and ensures that the information is safe and hidden.

References

[1] Thompson K, Miller G J, Wilder R. Wide-area Internet traffic patterns and characteristics[J]. IEEE network, 1997, 11(6): 10-23.
[2] Lin C C, Tsai W H. Secret image sharing with steganography and authentication[J]. Journal of Systems and software, 2004, 73(3): 405-414.
[3] Clause J, Orso A. Camouflage: automated anonymization of field data[C]//Proceedings of the 33rd International Conference on Software Engineering. ACM, 2011: 21-30. [3] A. Name, "Dissertation Title", M.S.(or Ph.D.) thesis, Department, University, City, Country, Year.
[4] Anitole G. Method for developing natural camouflage patterns: U.S. Patent 4,576,904[P]. 1986-3-18.
[5] Habes A. Information hiding in BMP image implementation, analysis and evaluation[J]. Saint Petersburg Institute for Informatics and Automation, Russian Academy of Sciences, Saint Petersburg, Russia Received February, 2006, 26.
[6] Lu H, Wan B. Information Hiding Algorithm Using BMP Image[J]. Journal of Wuhan University of Technology, 2006, 28(6): 96-98.
[7] Bourke P. BMP image format[J]. BMP Files. July, 1998.
[8] Konstantinides K, Bhaskaran V, Beretta G. Image sharpening in the JPEG domain[J]. IEEE transactions on image processing, 1999, 8(6): 874-878.
[9] Yukihiro A, Takeshi A, Nakajima M. A fast DCT-SQ scheme for images[J]. IEICE TRANSACTIONS (1976-1990), 1988, 71(11): 1095-1097.
[10] Hamilton E. JPEG file interchange format[J]. C-Cube Microsystems, 1992.

Yu Tang was born in 1992, currently studying at the Jiangsu University of Science and Technology. She majored in control engineering and her research direction is information security.