Steganography Technique to Hide a Secret Message in QR Code

A. Elhawil
Computer Engineering Department,
Faculty of Engineering,
University of Tripoli
Tripoli – Libya
A.elhawil@uot.edu.ly

A. Alsellami
Computer Engineering Department,
Faculty of Engineering,
University of Tripoli
Tripoli, Libya

S. A. Talha
College of Electronic Technology
Tripoli, Libya
saseya@ltt.net

I. Mohamed
Computer Engineering Department,
Faculty of Engineering,
University of Tripoli
Tripoli, Libya

R. Awein
Computer Engineering Department,
Faculty of Engineering,
University of Tripoli
Tripoli, Libya

Abstract— **This paper proposes a new method to hide a message in Quick Response code (QR code). This type of image steganography is quite a challenge because the secrete message should not be detected by the standard QR code reader. In addition, it has to be stored in an appropriate location that does not affect the original data of the QR code. All these requirements are studied. The analysis method is presented and discussed.**

Index Terms:**Image stenography, QR code, encoding, decoding.**

I. INTRODUCTION

Quick Response code (QR code) is a type of two-dimensional bar code. It is a machine-readable optical label contains information about the item to which it is attached [1]. Although the QR code system was initially invented in 1994 by Denso Wave to track parts in vehicles manufacturing; it is now broader used in various applications such as commercial tracking applications and convenience-oriented applications aimed at mobile-phone users (termed mobile tagging) [2].

QR code contains information in both the horizontal and vertical direction, they met the need for high data density and small size. QR codes may be used to display text, to add a vCard contact to the user's device, to open a Uniform Resource Identifier (URI), or to compose an e-mail or text message. Nowadays, users can generate and print their own QR codes using free available software applications.

Steganography is a method of embedding a message in a cover utterance for secure communication. There are different types of steganography, image steganography audio steganography, video steganography and network protocol steganography. Each method has its own features [3].

Since QR codes have a large storage capacity, malicious message content that could include terroristic plans, attack coordinates and links to landmark attractions, etc. can be stored.

Received 14 November, 2017; revised 25 December, 2017; accepted 29 December, 2017.

Available online Januarv 1. 2018.

The aim of this paper is to develop a method to embed a secret message in a QR code without effecting the original QR code content. The treated QR code can be scanned using any standard QR reader, but only our developed software is able to retrieve the secret message. Moreover, the security level is increased by encrypting the text before the embedding process.

Furthermore, only authorized users, who have the software application and the encryption algorithm keys, are able to extract and decrypt the secret message from the published QR code.

II. QR CODE STEGRAOGRAPHY

Recently QR code steganography has become a new research issue. Many papers have been published that deal with how to exchange encrypted message in QR code [4] and [5].

Being able to hide secret messages within general QR code symbols creates endless possibilities for discreet communication through QR codes [4].

In cryptography, the encryption is the process of converting a messages (or information) in such a way that eavesdroppers or hackers cannot read it, but that authorized parties can. In an encryption scheme, the message or information (referred to as plaintext) is encrypted using an encryption algorithm, turning it into an unreadable cipher text (ibid.) [6]. This is usually done with the use of an encryption key, which specifies how the message is to be encrypted, any adversary that can see the cipher text should not be able to determine anything about the original message [7].

Only the authorized party is able to decrypt the cipher text using a decryption algorithm, that usually requires a secret decryption key, that adversaries do not have access to. For technical reasons, an encryption scheme usually needs a key generation algorithm to randomly produce keys [8].

The most important problem in this research is how to embed the secret message in the QR code without affecting the original message. In addition, the message must be secure and unable to be hacked. The most common QR code manipulation technique has been to embed QR code symbols with phishing links to websites containing viruses. By using this technique an attacker

can utilize non-human dabble symbols to gain sensitive information or monetary benefit. A far greater threat exists in the idea proposed in this paper, a terrorist or radical could utilize QR codes to hide secret information which can only be read by privileged members of the organization [4].

III. QR CODE STRUCTURE

Each QR code is a regular square array constructed of several nominally square modules, including an encoding region and function patterns, namely finder, separator, timing patterns, and alignment patterns [9], as shown in Figure 1. Function patterns cannot be used to encode data, and the QR code symbol is surrounded by quiet zone on all four sides, as shown in Figure 1.

There is a total of 40 versions of the QR code, from 21 by 21 (version 1) modules to 177 by 177 (version 40) modules, increasing in steps of 4 modules per side, Naturally, higher versions are used to encode larger amounts of data, as shown in Figure 2.

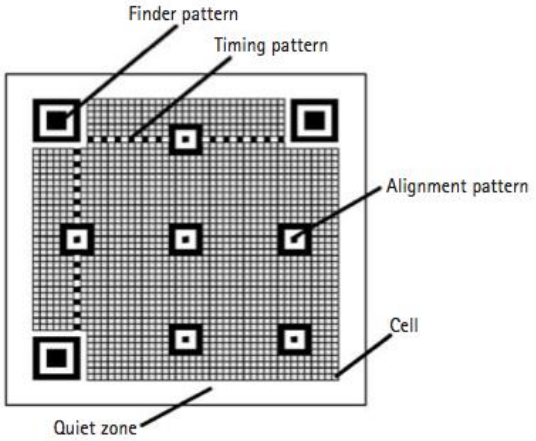

Figure 1. Structure of QR Code

The QR code consists of the following main parts [10]:

1. *Alignment pattern*: the alignment pattern is only included in the rendered QR code in version 2 and above. Its purpose is to allow the decoder to scan a skewed image, and convert it to the virtual grid of black and white modules, representing the encoded data [2]. The alignment pattern is made of concentric squares, much like the finder patterns, with the center being a single black module.
2. *Finder pattern:* consists of three identical structures that are located in all corners of the QR code except the bottom right corner. Each pattern is based on a 3x3 matrix of black modules surrounded by white modules that are again surrounded by black modules. The Finder Patterns enable the decoder software to recognize the QR code and determine the correct orientation.
3. *Timing pattern*: this pattern is an alternating stripe of black and white modules, starting at the lower left corner of the upper right Finder Pattern, going horizontally to the upper left finder pattern and then going vertically to the lower left finder pattern.
4. *Format information*: the format data is information, pertaining to the Masking rule used in the QR code, along with error correction level. When the data in the QR code is encoded, some of the modules are inverted, in accordance with a predefined rule, in order to improve readability, and ensure that there are no big clusters of same-colored modules. This process is called masking, and the masking information is included in the format data, to alert the decoder that certain modules have been inverted. The format data is encoded in 15 bits. One full copy of the format data is located around the upper left finder pattern. A second copy, divided into 7 and 8 bits, is located next to the other two finder patterns.
5. *Version information*: the version data includes information on which version the QR code is. This data is encoded into 18 modules, in a 6 by 3 matrix. Two copies of the version data matrix are included in the QR code - one next to the upper right finder pattern, and the other next to the lower left one.
6. *Blank space*: additionally, around each QR code, there is an obligatory 4-modules-wide white space area.

Figure. 2 shows an example of how data is stored in the QR code. The number 14 in the image refers to the most significant bit of the format string, and the number 0 refers to the least significant bit [11].

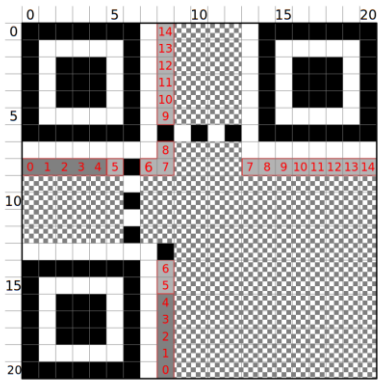

Figure 2. Locations of Data Stored in QR Code

IV. DATA SIZE OF QR CODE

The data consists of both the actual encoded data, entered by the user, and the error correction bits, calculated on that data [10]. QR code uses Reed Solomon algorithm to add error detection and correction information to source data. There are three main data types numeric (digits 0 - 9), alphanumeric (digits 0 – 9, upper case letters A – Z, nine other characters: space, $, %, *, +, -, / and :) and binary data. Each QR code symbol version has a maximum data capacity according to the

amount of data, character type and error correction level. In other words, as the amount of data increases, more modules are required to comprise QR code, resulting in larger QR code symbols. The smallest module is 21 x 21 and the largest module is 177 x 177. Table 1 lists the data capacity of some QR code versions.

Table 1. OR Code Data Capacity

Versin	Modules	Error Correction	Numerc only	Alphanumec
1	21x21	L	41	25
		M	34	20
		Q	27	16
		H	17	10
2	25x25	L	77	47
		M	63	38
		Q	48	29
		H	34	20
3	29x29	L	127	77
		M	101	61
		Q	77	47
		H	58	35
4	33x33	L	187	114
		M	149	90
		Q	111	67
		H	82	50
14	73x73	L	1,101	667
		M	871	528
		Q	621	376
		H	468	283
40	177x177	L	7,089	4,296
		M	5,596	3,391
		Q	3,993	2,420
		H	3,057	1,852

V. QR CODE ERROR CORRECTION LEVEL

In case of QR code is damaged, the stored data can be read as much as 30% of the code is corrupted. QR code uses Reed-Solomon Error Correction algorithm. The error correction feature is implemented by adding a Reed-Solomon Code to the original data. There are four levels of error correction, each one adds different amounts of "backup" data depending on how much damage the QR code is expected to suffer in its intended environment, and hence how much error correction may be required [12]:

1. Low Level (L): approximately 7% of code-words can be restored.
2. Medium Level (M): it allows restore 15% of code-words.
3. Quality level (Q): approximately 25% of code-words can be restored.
4. High level (H): it allows recovery of up to 30% data loss.

VI. PROPOSED SENARIO

As mentioned in the previous section, the QR code with high correction level is able to recover 30% of full size of QR code, and this error level is stored in the middle of the QR code. In this paper, another QR code smaller than the original one is generated. This small QR code contains the secret data to be stored in the original one. The following steps illustrate the process:

1. Let the size of the secret message we need to store is 100 character of alphanumeric data type.
2. Based on the data size, which is 100 characters, and from Table 1, the most appropriate version of QR code is version 4. That is because the capacity of this version is 114 bytes with the low error correction level. So, the secret message is first stored in QR code version 4.

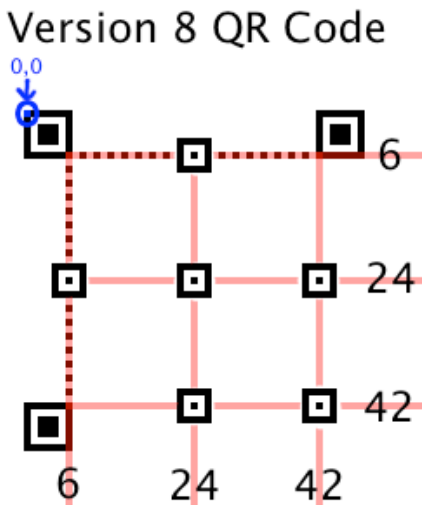

Figure 3. Locations of Alignment Patterns in Version 8 [11]

3. Now, it is required to embed this small QR code in a bigger one. Based on the version of the small QR code, that contains the secret message, a bigger version of QR code is chosen. The most suitable version for this purpose is version 14. This version where is chosen based on several tests done using Image Editor (Photoshop) and standard QR code reader application. In each test, the secret QR code is placed in different versions of the original QR code in a manner that, the original data is not affected. From QR code structure, most of the patterns of QR code are stored in the corners, as shown in Figure. 1, with the exception of the alignment pattern. The position of the alignment pattern varies depending on the QR code version. For example, the alignment pattern locations of version 8 is 6, 24, and 42. All combinations of these three numbers are used as coordinates for alignment patterns. Figure. 3 shows the locations [11]. By studying a set of versions, we found out that the center of version 14 does not contain any information. That is because its alignment pattern locations are 6, 26, 46 and 66. By doing simple calculations the center of the 73 x 73 square is empty and can fit QR code of version 4.

However, when we place the secret QR code (version 4) in the middle of the big QR code (version 14), we have noted that the secret QR code could be clearly identified as shown in Figure. 4. To solve this problem, we removed the finder patterns of the secret QR code before embedding. The finder patterns help the QR code reader to determine the correct orientation. These identical

patterns however, can be removed and regenerated in the extraction process. Figure. 5 depicts QR code of version 4 after removing the finder patterns. Now the secret QR code ready to be embedded in version 14. Figure. 6 shows the final QR code. It looks as a normal QR code and we have tests it using different standard QR code readers. All the QR readers could recognize the original message only. Our secret QR code could not be detected.

Figure 4. a) Original QR code. b) Combined QR Codes

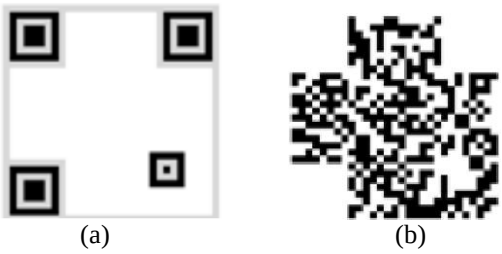

Figure 5. a) Finding Patterns of QR Code Version 4 QR Code. b) QR Code after Removing the Patterns

In order to read the secret information from small QR code, we first extract the secret QR (Version 4) From big QR code (Version 14) because we know exactly where is it stored, then we return the finder patterns that we removed before.

Figure 6. The Resultant QR Code.

The last problem that we have is what if someone has any doubts about our QR code for any reason. To avoid this, and also to increase the level of the security, the secret message is encrypted and decrypted using utf-8 standard. Both encryption and decryption processes request a secret key. This key is only known by people authorized to read secret data

VII. CONCLUSION

In conclusion, this paper proposes a method to hide a secrete data in QR code. The resultant QR code, which contains both the original information and secret message, can be read by any standard QR reader. That means the original information of QR code is not affected. In addition, the secret message cannot be detected by the standard QR reader. The receiver should use our application to decode the secret message. The proposed method is performed on QR code version 4 and 14. The results are very motivating.

REFERENCES

[1] Handbook of Augmented Reality, 1st ed., Springer Science & Business Media, pp. 341, 2011.

[2] *QR Code*. Wikimedia Foundation, Inc.. 2017. Available: https://en.wikipedia.org/wiki/QR_code#cite_note-About2DCode-4/

[3] H. Singh, "*Analysis of Different Types of Steganography*, International Journal of Scientific Research in Science, Engineering and Technology IJSRSET, vol. 2, Issue 3, pp. 578-582, June, 2016.

[4] D. J. Ohana, and N. Shashidhar , "*QR Code Steganography*, " presented at: International Conference on Security and Management (SAM); 2013; Las Vegas.

[5] J. Rouillard, "*Contextual QR code, Computing in the Global Information Technology*, " pp. 50-55, Aug. 2008.

[6] Leighton Johnson, Conducting a Successful Incident Response: Computer Incident Response and Forensics Team Management, 2013, p10.

[7] V. Gupta and S. Sharma, "*Encryption and decryption using one pad time algorithm in mac layer*, " International Journal of Innovative Research in Science, Engineering and Technology Vol. 2, Issue 6, June 2013

[8] P. Gaur1, P. Singh, *"A Review of Geometry Based Symmetric Key Encryption Using Ellipse*, " International Journal of Computer Science and Mobile Computing, IJCSMC, vol. 2, Issue. 6, June 2013, pp.1 – 6.

[9] KarelBRG, "*QR Code Introduction*, " https://www.scribd.com/document/152268386/QR-Code-Introduction.

[10] Telerik UI for WindowsPhone Visual Structure, http://www.telerik.com/help/windows-phone/radbarcodeqr-qrcode-visual-structure.html, 2016.

[11] "*2017's Top 6 Label Makers*, " http://www.thonky.com/qr-code-tutorial/format-version-information

[12] "*QR Code Error Correction*,", https://blog.qrstuff.com/2011/12/14/qr-code-error-correction, April, 2017.

BIOGRAPHIES

Amna Elhawil received BSc degree in the computer engineering from University of Tripoli (Tripoli -Libya). She got Msc degree in 2006 and PhD degree in 2010 from the Department of Electronics and Information Processing (ETRO) of Vrije Universiteit Brussel (VUB) - Belgium. She is currently assistance professor in the Department of Computer Engineering at University of Tripoli /Libya.

Ashor Alsellami is an assistant professor at Computer Engineering Department Faculty of Engineering, University of Tripoli (Tripoli - Libya). He did his postgraduate studies at Manchester university and also Istanbul technical university. He is interested in coding theory, signal processing and channel modeling.

Saad A. Talhah. was born in Tripoli-Libya, on Feb. 25 1961. He received Bsc degree in Electrical and Electronics from Higher Institute of Electronics Ben-Walid Libya 1982., He got Msc degree in Electrical and Electronics System Engineering from (UKM) Malaysia 1995. Moreover, he got PhD degree in Electronic engineering from University of Leeds UK 2005. Where he is currently associate professor at College of Electronic Technology-Tripoli Libya and in the same time Dean of the College. His research field in cryptography

Ibrahim Mohamed received the BSc. degree in computer engineering in 2015 from University of Tripoli (Tripoli -Libya).

Ramy Awein received the BSc. degree in computer engineering in 2015 from University of Tripoli (Tripoli -Libya).