Journal of Cyber Security
DOI:10.32604/jcs.2021.026022
Article

Tech Science Press

A Distributed Secret Sharing Method with QR Code Based on Information Hiding

Pengcheng Jiang and Yu Xue[*]

Nanjing University of Information Science & Technology, Nanjing, 210044, China
[*]Corresponding Author: Yu Xue. Email: xueyu@nuist.edu.cn

Received: 13 December 2021; Accepted: 21 December 2021

Abstract: QR codes are applied widely on the Internet and mobile devices in recent years. Its open standards and the characteristics of easy to generate lead to anyone can generate their QR code easily. Also, the QR code does not have the ability of hiding information, which leads to everyone can get access to read the content in any QR code, including those hiding some secret content in the bytes of QR code. Therefore, in recent years, information tampering and information leakage cases caused by poor security of two-dimensional code occur frequently, especially in the financial field and multi-party verification scenarios. QR codes are almost impossible to use in these scenarios. Therefore, this paper proposes a distributed information sharing method based on information hiding QR code. This method can make secret code in QR code safer and robust, and the secret shared between receivers can be used for decryption and attacking detection. Therefore, on the one hand, the information hiding method can maximize the capacity of embedded secret information, on the other hand, it can prevent attacks by disguised attackers and recover hidden secret information through reconstruction. This paper illustrates the feasibility of this scheme through the form of theoretical proof.

Keywords: QR code; image encryption; information hiding; secret sharing

1 Introduction

Two-dimensional code is formed based on of one-dimensional code in the vertical direction of the expansion of the data dimension. There are many types of two-dimensional codes, among which the most common is now widely used quick response code (referred to as QR Code). Compared with other types of bar codes, QR codes can store more information, and through the scanner to jump or respond to different protocols can achieve a lot of convenient operations, such as quick access to the URL address, quick access to business card information, speed dial, fast connection to WLAN, etc. Because QR code belongs to open standards, so almost everyone can use tools to generate their QR code, which creates a forge the QR code is rampant, has great potential safety hazard, and because the QR codes Bunsen do not have the function of hidden information, and no design protection mechanisms, so any device can read the same information, it includes both private and confidential information and is therefore not secure. This paper will introduce a method of information hiding based on QR code, embed ciphertext into the QR code without changing the original information of the QR code, to form a special QR code that only some people can read specific information. In some scenarios (such as electronic coupons, authorization in e-commerce, etc.), if two-dimensional code is used as the carrier of information interaction and security can be guaranteed in the process, the system will be easier to use.

2 Related Technologies and Algorithms

In this section, we will introduce the related work, including the principle of QR codes, the distributed technology, and the digital watermarking technology.

2.1 QR Code

Two-dimensional code (2-dimensional Bar Code) has been widely used on the Internet and mobile devices in recent years. QR code (quick response code) is a widely used two-dimensional matrix representation, developed by Denso Wave in Japan in 1994. It is a black and white graphic used to record data symbol information distributed on a plane with a specific geometric figure according to a certain rule. It is mainly composed of four functional modules (see Fig. 1): detection graph, positioning graph, data area, and correction graph.

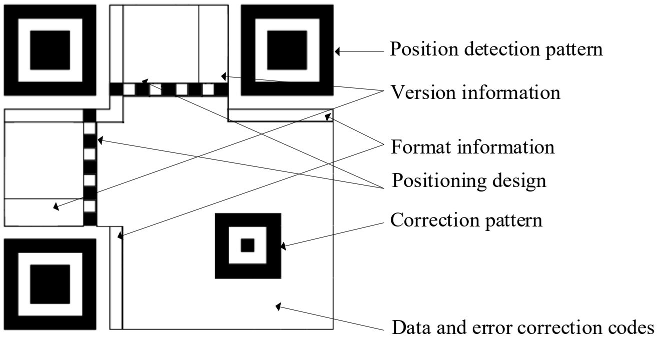

Figure 1: Basic structure of two-dimensional code

QR code coding process is roughly divided into data analysis, data coding, error correction coding, data organization, application data mask, filling format, and version information. The two-dimensional code standard provides 40 two-dimensional code versions, the larger the version can fill in the more data. Another feature of QR code is its reliability, QR code can be set error correction level, divided into L (Low), M (Medium), Q (Quartile), and H (High), respectively 7%, 15%, 25%, 30% error correction rate (as shown in Table 1).

Table 1: Error correction level of QR code

Error correction level	Max rate to correction (%)
L (Low)	7
M (Medium)	15
Q (Quartile)	25
H (High)	30

2.2 Distributed Technology

Distributed computing is a type of computer science that studies how to divide a large problem into many small parts, then distribute these parts to multiple terminals for processing, and finally combine the results of these calculations to get the final result. Distributed computing makes it possible for data to be distributed on different devices, which enables the assumption that only multiple terminals are needed to achieve the complete function, eliminating the response design of the server.

2.3 Digital Watermarking Technology

Digital watermarking refers to embedding identification information into carriers, such as images, videos, audio, documents, and software, or indirectly expressing identification information by modifying the structure of a specific area. These operations do not affect the normal use of the original carrier, so this technology is often used for copyright protection, anti-counterfeiting, confidentiality checks, and so on.

3 Methodology

This section describes the approach used in this article.

3.1 Ciphertext Embedding Based on Information Hiding

At present, most QR codes realize ciphertext sharing by preparing a database server, and the terminal will scan the QR code and pass the parameters to the server for interaction [1], to obtain the ciphertext data corresponding to the content of the QR code. In this way, everyone can easily obtain the complete information in the two-dimensional code, lack of security in the process of sharing.

To solve this security problem, the most common method is to embed a certain amount of ciphertext information in the QR code to achieve information hiding. In the selection of information hiding mode and location, there are the following two commonly used schemes: 1) embedded digital watermark into the image; 2) directly embedded in the two-dimensional code binary data.

Scheme 1 uses the scheme [2–4] of embedding digital watermark directly into the image. DWT, DCT, and DFT are used to embed the watermark into the two-dimensional code image domain. The capacity of the watermark depends on the size of the image and has nothing to do with the size of the two-dimensional code. However, when reading the two-dimensional code, it is necessary to carry out bilinear interpolation transformation, morphological repair, and Bose-Chaudhuri-Hocquengham (BCH) code error correction in order to extract the watermark from the distorted two-dimensional code when the pixel distortion is caused by factors such as the recognition Angle and the quality of the image itself. In addition, the method of scaling grid points [5–6] is used to appropriately enlarge or shrink some grid points, for example, the scaling represents 0 and 1, respectively, so as to realize information hiding and embedding. Then Gao and Sun directly adjusted the width of the row and column of the two-dimensional code module [7] and embed the watermark into the two-dimensional code label. The principle of the two implementations is that the QR code automatic recognition allows a certain percentage of grid error, when reading can be read normally, through some operations can also get the information. However, the acceptable range of grid point size adjustment is limited, and the capacity of embedded data of two-dimensional code is directly related to the size of two-dimensional code, so the effective capacity of embedded watermark is smaller than the scheme [2–4].

The second approach is to directly embed the information into the two-dimensional code of the binary information area. The basic idea of this scheme is to use the code and error correction mechanism of two-dimensional code, modify the lattice value of some two-dimensional code or directly modify the code value of two-dimensional code will not affect the normal reading of two-dimensional code, and can realize information hiding. The method adopted by Tang [6] selects the position of random sequence in the R-S coding process of two-dimensional code for data embedding and extracts information by constructing the same logical sequence in the embedding process during reading. In addition, Lin [8] embedded information through error-correcting code words.

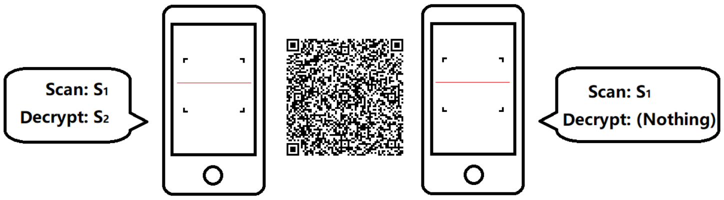

Figure 2: QR code information hiding result

In this paper, an adaptive ciphertext embedding method is proposed to improve the two traditional schemes. After analysis, in some cases, the simple embedding of two-dimensional code data words or the embedding of error correction code is easy to cause the waste of data area, only a reasonable allocation of two-dimensional code embedded area, in order to minimize the two-dimensional code size, have a greater hiding capacity. The two-dimensional code processed in this way can still obtain the plaintext data of the two-dimensional code with other scanners, but the hidden data hidden in the two-dimensional code can be obtained with specific software or scanners (as shown in Fig. 2). The left phone using the method proposed in this paper can decrypt the secret information and the right phone using the traditional method cannot.

The specific implementation steps are as follows:

1) Analyze the size of the hidden data volume, and select the appropriate version and error correction level based on the size of the original information of the two-dimensional code. The number of redundant code words and error correction code words in the data code word of the two-dimensional code is directly related to the size of the data volume that the two-dimensional code can hide (as shown in Table 2). When the total amount of data is small, a smaller version or a smaller error correction level can be used to generate a smaller size of the two-dimensional code. When the amount of hidden data is small but the amount of original data is large, the automatically determined version and the smaller error correction level can be adopted. When the amount of hidden data is no longer significantly smaller than the amount of original data, the larger error correction level or even the two-dimensional code version larger than the automatically determined version should be selected to obtain more embedding space.

Table 2: Max & min capacity of QR code

Version	Error correct level	Bytes for correction	Blocks for collection	Bytes in each block	Bytes for information	Bits for information
1	L	7	1	19	19	152
	M	10	1	16	16	128
	Q	13	1	13	13	104
	H	17	1	9	9	72
20	L	224	3	107	861	6888
			5	108		
	M	416	3	41	669	5352
			13	42		
	Q	600	15	24	485	3880
			5	25		
	H	700	15	15	385	3080
			10	16		
40	L	750	19	118	2956	23648
			6	119		
	M	1372	18	47	2334	18672
			31	48		
	Q	2040	34	24	1666	13328
			34	25		
	H	2340	20	15	1276	10208
			61	16		

2) small range hidden data will be hidden information is converted to a bit stream, using data from a QR code directly on the specific area code word for exclusive or operation, use the QR code itself, to correct the error correction capability of the modified blocks in, thus will not affect the reading, the filling capacity can use the formula $C = \lfloor E/2 \rfloor \times 8$ get (C for filling capacity, E is the number of error correcting code words of the selected version). During scanning, the hidden information is extracted by

XOR operation with the QR code of the original data.

3) When more data needs to be hidden, the two-dimensional code of a higher version is selected, and the error correction level of a higher version is selected, and the hidden number is converted into bit stream. The hidden area should be planned according to the redundancy in the data area. When there is a large amount of redundancy in the data area, the redundant parts in the data area are preferentially selected for steganography (see Table 3 for the effect), and the contents of the redundant parts are separately identified and converted to secret information during reading. When the redundancy in the data area is small, error correction code words can be used to hide.

Table 3: The content before and after steganography in the two-dimensional code data area

Before hiding secrets	69 61 74 69 6f 6e 6a 6f 7b 6d 6e 24 86 66 72 2c 56 52 00 ec ec ec ec ec ec ec ec ec ec ec ec ec ec ec ec ec ec
After hiding secrets	69 61 74 69 6f 6e 6a 6f 7b 6d 6e 24 86 66 72 2c 56 52 00 63 a9 b4 6d 61 74 20 69 6e 66 6f 6e 20 61 ec ec ec ec

Compared with Lin's method of information hiding error correction codes, this method can have a greater data area utilization, especially when the two-dimensional code itself carries a small amount of data and the amount of hidden data is much larger than its own data, such as the two-dimensional code itself has only 20 bits of data, but there are 700 bits of hidden information. According to the calculation results of scheme [9], the two-dimensional code of 20-L will be selected to hide the data, but in fact, the two-dimensional code of 20-L can store far more data than the 20 bits (see Table 2). This means that there will be a lot of redundancy in the data code part. According to this scheme, the two-dimensional code of a slightly lower version and the error correction level of H level can be selected to write the redundant part in the data code word.

3.2 Secret Distributed Sharing Methods

The concept of secret sharing was first proposed by Shamir [10] and Blakley [11] based on Lagrange interpolation polynomials and hyperplane geometry respectively. Secret sharing schemes generally include secret allocation algorithm and secret reconstruction algorithm. When the allocation algorithm is implemented, the secret publisher divides the shared secret into several parts and distributes them to the participants through certain policies. Each participant gets a sub-secret of a secret. During secret reconstruction, all or a specified number of participants will constitute an authorization set, and the secret can be recovered by secret reconstruction algorithm when the number of participants in the authorization set is secret. Any participant set with less than the number of participants cannot calculate the secret. In the $(t; n)$ threshold scheme, T or more participants can form an authorized set to perform the secret reconstruction algorithm to obtain the secret. This means that when the participants meet this condition with dishonest participants still can restore the secret, even if the attack was detected, but the reconstruction algorithm for the dishonest participants can still execute and can successfully get a secret, for the other participants of the honesty, reconstruction algorithm can lead to cannot calculate the result or calculate the wrong results. Tompa et al. proposed to hide secrets in a set of sequences [11]. Laih et al. proposed v-fair $(t; n)$ secret sharing scheme improvement [12]. In 2013, Tian et al. [13] used the consistency of sub-secrets to detect the existence of chelators. However, Harn later pointed out in literature [14] that Tian's fair scheme [13] could not completely resist asynchronous attacks, and constructed a method $(t; n)$ Threshold scheme [15]. Only in the $(n; n)$ secret sharing scheme can all participants constitute an authorization set, so that the final secret can be obtained through the reconstruction algorithm. Taking the algorithm put forward by Li et al. [16] as an example, the computation of the n messages depends on secrets itself, and only inverse operation is able to reconstruct the ciphertext, but this method does not take into account not honest attack once an attacker is involved. The secret will not be able to reconstruct when the honest participant number is n, and the attacker can even to refactor secret and honest but not directly.

Based on the above problems, this paper proposes a new secret sharing scheme. In the secret allocation stage, a strategy similar to (n, n) secret sharing scheme is used to divide secret S into m sub-secrets $(S_1, S_2 \cdots S_m)$, and form n messages $(M_1, M_2 \cdots M_n)$ and the validation matrix, and finally the message M_i is assigned to the participant along with the validation matrix. Before the secret reconstruction, m participants are successively selected to verify the security. The secret reconstruction can be carried out only after all participants have passed the verification. Therefore, the scheme proposed in this paper will eliminate the attacker before all the sub-secrets meet, and the secret can be reconstructed only when all the participants participate. The specific scheme includes two sub-schemes, secret allocation, and secret reconstruction, as follows:

Secret distribution scheme:

1) The secret publisher randomly generates $(m-1)$ keys, denoted as $S_1, S_2 \ldots S_{\mathrm{m}-1}$;

2) Calculate S_m, and the calculation formula is as following:

$$S_m = S_1 \oplus S_2 \oplus \cdots \oplus S_{m-1} \oplus S$$

3) Each key $(S_1, S_2 \cdots S_m)$ is divided into $(n+1)$ parts, for example, S_i is divided into $(n+1)$ parts $S_{i0}, S_{i1}, S_{i2} \ldots S_{in}$, S_{ij} represents the j^{th} interval secretly divided into the i^{th} submitter, and satisfies the following formula:

$$S_i = [S_{i0}\ S_{i1}\ S_{i2} \cdots \cdots S_{in}]$$

4) Put S_{ij} (1 ≤ i ≤ m， 1 ≤ j ≤ n) into the matrix $A_{m\times n}$, transpose the matrix, and put S_{i0} (1 ≤ i ≤ m) into the matrix $B_{n\times 1}$, the following are the two matrices, and record their corresponding relations:

$$\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ S_{m1} & S_{m2} & \cdots & S_{mn} \end{bmatrix}^T_{m\times n} \qquad \begin{bmatrix} S_{10} \\ S_{20} \\ \vdots \\ S_{m0} \end{bmatrix}_{m\times 1}$$

5) Multiply the transposed matrix with the matrix $B_{n\times 1}$, and the result is $R_{n\times 1}$:

$$\begin{bmatrix} S_{11} & S_{21} & \cdots & S_{m1} \\ S_{12} & S_{22} & \cdots & S_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ S_{1n} & S_{2n} & \cdots & S_{mn} \end{bmatrix}_{n\times m} \cdot \begin{bmatrix} S_{10} \\ S_{20} \\ \vdots \\ S_{m0} \end{bmatrix}_{m\times 1} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix}_{n\times 1}$$

6) The combination matrices $A^T_{m\times n}$ and $R_{n\times 1}$ form the following matrix $M_{n\times(m+1)}$:

$$\begin{bmatrix} S_{11} & S_{21} & \cdots & S_{m1} & R_1 \\ S_{12} & S_{22} & \cdots & S_{m2} & R_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{1n} & S_{2n} & \cdots & S_{mn} & R_n \end{bmatrix}_{n\times(m+1)}$$

7) Each row of the matrix $M_{n\times(m+1)}$ is assigned to the participant as message M_{i} along with the matrix $B_{n\times 1}$, and the allocation algorithm is complete.

Covert reconstruction scheme

1) Random negotiation among n participants results in m participants exchanging row matrices obtained by the matrix $M_{n\times(m+1)}$;

2) Each participant is segmented and pieced together to form matrices A'^T_{m*n} and R'_{n*1} (as is shown in the following formula, S'_{ij} represents the j^{th} interval divided by the i^{th} user's sub-secret):

$$\begin{bmatrix} S'_{11} & S'_{21} & \cdots & S'_{m1} \\ S'_{12} & S'_{22} & \cdots & S'_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ S'_{1m} & S'_{2m} & \cdots & S'_{mm} \end{bmatrix}_{m\times m} \qquad \begin{bmatrix} R'_1 \\ R'_2 \\ \vdots \\ R'_m \end{bmatrix}_{m\times 1}$$

3) Construct the following equation and solve it:

$$\begin{bmatrix} S'_{11} & S'_{21} & \cdots & S'_{m1} \\ S'_{12} & S'_{22} & \cdots & S'_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ S'_{1m} & S'_{2m} & \cdots & S'_{mm} \end{bmatrix}_{m\times m} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix}_{m\times 1} = \begin{bmatrix} R'_1 \\ R'_2 \\ \vdots \\ R'_m \end{bmatrix}_{m\times 1}$$

4) The obtained column matrix is compared with matrix $B_{n\times 1}$. If it is consistent, it indicates that m participants participating in this calculation are honest participants; if not, it indicates that there are attackers in each participant participating in this calculation;

5) Repeat Steps (1) to (4) to identify all participants through a combination of honest participants and unauthenticated participants;

6) If n participants are all honest participants, perform the final reconstruction calculation: n participants exchange the row matrix obtained by the matrix $M_{n\times(m+1)}$, and then remove the last column and transpose to get $A'_{m\times n}$:

$$\begin{bmatrix} S'_{11} & S'_{12} & \cdots & S'_{1n} \\ S'_{21} & S'_{22} & \cdots & S'_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ S'_{m1} & S'_{m2} & \cdots & S'_{mn} \end{bmatrix}_{m\times n}$$

7) Use the corresponding relation to form, corresponding to the row matrix, form sub-secret S'_1, $S'_2 \cdots\cdots S'_{\mathrm{m-1}}$;

8) Apply the following formula to obtain the secret:

$$S = S'_1 \oplus S'_2 \oplus \cdots\cdots \oplus S'_m$$

3.3 QR Distributed Secret Sharing Method Based on Information Hiding

The above two-dimensional code information hiding scheme is combined with the distributed secret sharing scheme, and the messages required by each participant are hidden in the two-dimensional code as a carrier, which can only be recognized by specific devices, or parameters are added in the process of two-dimensional code steganography, so that only specific users can obtain the messages. The specific plan is as follows:

Distribution process (as shown in Fig. 3):

1) Perform the above secret allocation algorithm to obtain n messages required by users;

2) Form a standard two-dimensional code, the content of the two-dimensional code can be the download address of the scanner, promotion page, etc.;

3) Apply the above embedding process to embed the message of each participant into the QR code and distribute the QR code.

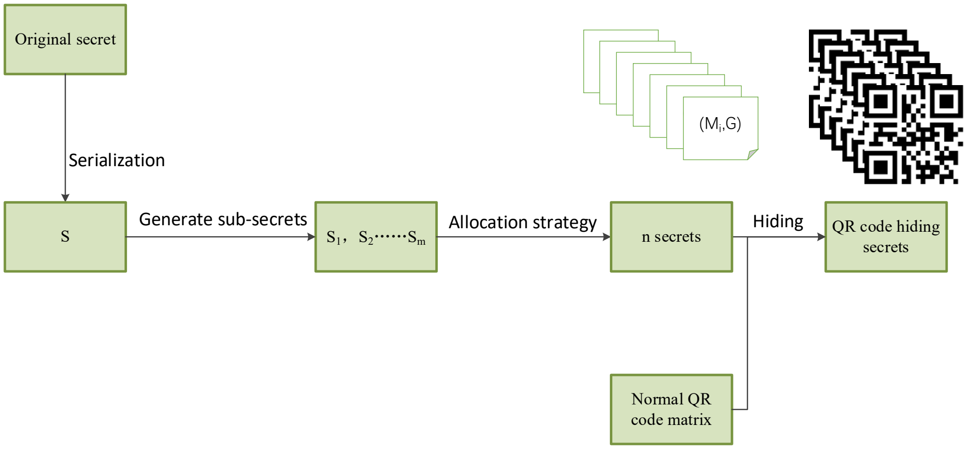

Figure 3: Distribution process

The refactoring process (Fig. 4):

1) The hidden information of the two-dimensional code can be obtained by the scanner of the participant (M_i, G);

2) Perform the above secret reconstruction algorithm to recover the information after eliminating the attacker.

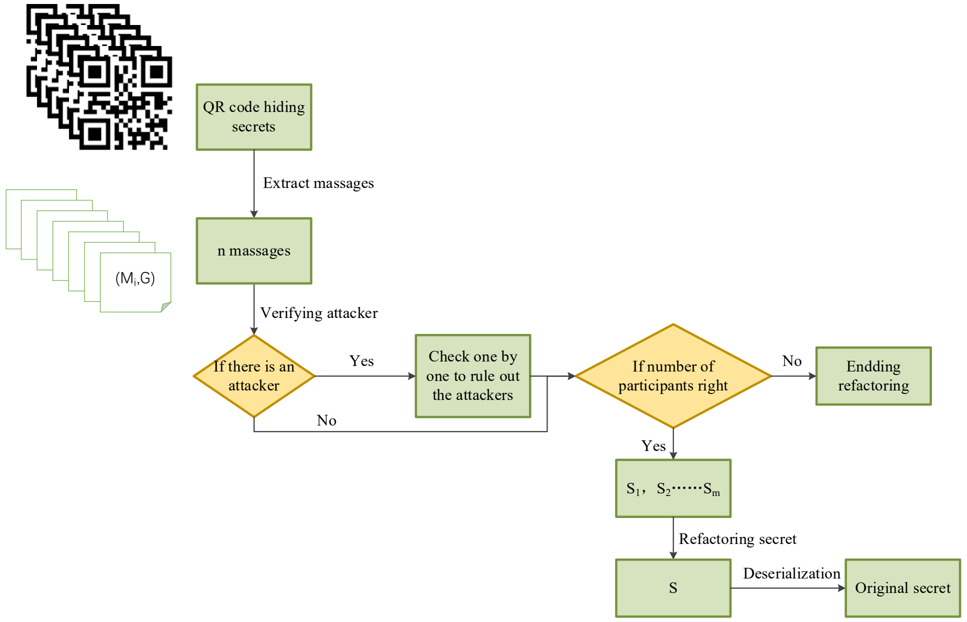

Figure 4: Refactoring process

4 Conclusion

This paper uses two-dimensional code to hide information, which is different from the traditional scheme of using pictures to hide information. This paper relies on the principle of two-dimensional code design and error correction mechanism and obtains hidden information through XOR operation after error correction. The distributed key sharing design proposed in this paper can be checked before secret reconstruction to exclude attackers. Combined with two-dimensional code information hiding technology, it will meet many specific requirements, such as multi-party authorization of e-commerce and identification of electronic coupons. It is worth mentioning that the secret sharing scheme proposed in this paper can be combined with other information hiding schemes, and the specific hiding schemes can be adjusted according to the actual needs.

Funding Statement: This work was partially supported by the National Natural Science Foundation of China (61876089, 61876185, 61902281, 61375121), the Opening Project of Jiangsu Key Laboratory of Data Science and Smart Software (No. 2019DS301), the Engineering Research Center of Digital Forensics, Ministry of Education, the Science and Technology Program of Jiangsu Province Construction System (2020JH08), and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. C. Chuang, Y. C. Hu and H. J. Ko, "A novel secret sharing technique using QR code," *International Journal of Image Processing*, vol. 4, no. 5, pp. 468–475, 2010.

[2] M. Sun, J. Si and S. Zhang, "Research on embedding and extracting methods for digital watermarks applied to QR code images," *New Zealand Journal of Agricultural Research*, vol. 50, no. 5, pp. 861–867, 2007.

[3] L. Li, R. L. Wang and C. C. Chang, "A digital watermark algorithm for QR code," *International Journal of Intelligent Information Processing*, vol. 2, no. 2, pp. 29–36, 2011.

[4] S. Rungraungsilp, M. Ketcham, V. Kosolvijak and S. Vongpradhip, "Data hiding method for QR code based on watermark by compare DCT with DFT domain," in *Proc. ICCCT*, Allahabad, Uttar Pradesh, India, pp. 144–148, 2012.

[5] P. Y. Lin and Y. H. Chen, "High payload secret hiding technology for QR codes," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, pp. 1–14, 2017.

[6] L. Tang, "The research and implementation on two-dimensional barcode information hiding," M.S. dissertation, Hunan University, China, 2013.

[7] M. Gao and B. Sun, "Blind watermark algorithm based on QR barcode," in *Proc. ISKE*, Shanghai, China, pp. 457–462, 2011.

[8] P. Y. Lin, "Distributed secret sharing approach with cheater prevention based on QR code," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 1, pp. 384–392, 2016.

[9] M. Tompa and H. Woll, "How to share a secret with cheaters," *Journal of Cryptology*, vol. 1, no. 3, pp. 133–138, 1989.

[10] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[11] G. R. Blakley, "Safeguarding cryptographic key," in *Int. Workshop on Managing Requirements Knowledge*, New York, NY, USA, pp. 313–317, 1979.

[12] C. S. Laih and Y. C. Lee, "V-fairness (*t, n*) secret sharing scheme," *IEE Proceedings–Computers and Digital Techniques*, vol. 144, no. 4, pp. 245–248, 1997.

[13] Y. Tian, J. Ma, C. Peng and Q. Jiang, "Fair (*t, n*) threshold secret sharing scheme," *IET Information Security*, vol. 7, no. 2, pp. 106–112, 2013.

[14] H. Lein, "Comments on "Fair (*t, n*) threshold secret sharing scheme," *IET Information Security*, vol. 8, no. 6, pp. 303–304, 2014.

[15] H. Lein, "Secure secret reconstruction and multi-secret sharing schemes with unconditional security," *Security and Communication Networks*, vol. 7, no. 3, 2014.

[16] Z. Li and G. Z. Cao, "Study on secret sharing scheme based on QR code," *Electronic Design Engineering*, vol. 26, no. 387, pp. 147–151, 2018.