偽装QRコードの構成とその効果，およびその対策について

大熊 浩也[1,a)]　瀧田　愼[1]　森井 昌克[1]

概要：二次元コードとしての QR コードはウェブページへのアクセス，特に最近では決済への利用等，幅広い用途に用いられている．QR コードは高い認識率を誇るものの，その内容を人が直接解釈できないことから，悪意のあるものが偽装した QR コードを作成し，不用意な操作から悪性サイトに導かれることが問題となっている．QR コードを作成するために，QR コードを作成できるシステム開発会社に依頼するか，もしくは QR コード作成ソフト，QR コード作成サイトを利用することが一般的である．しかし悪意を持ったシステム開発会社や作成サイトが偽装した QR コードを作成し，配布することも十分考えられる．偽装された QR コードは必ず悪性サイトに誘導されるがゆえに発見が容易であり，早い時点で対策が講じられる．著者らは既に誤り訂正符号の性質を用いて発見が困難な QR コードを開発している．前回の報告では，その具体的な作成方法において，一つのモジュールにドットを付加するまたは輝度を変えるなど，注意深く観察することによって通常の QR コードと識別できる例を与えた．偽装 QR コードは一つのモジュールだけでなく，QR コード全体に変更を加えることで，通常の QR コードとの識別が困難な偽装 QR コードを作成することも可能である．本稿では，そのような QR コードの作成方法を提案するとともに，偽装 QR コードの危険性を明示する．さらに，提案した偽装 QR コードに限らず，様々な方法で偽装される可能性がある QR コードについて，その対策を述べる．

キーワード：QR コード，偽装，誤り訂正符号，誤訂正

1. はじめに

二次元コードの一種である Quick Response（QR）コードは，スマートフォンに搭載されたカメラで容易に情報を読み取ることができるため，情報の伝達手段として幅広く使用されている．開発当初は QR コードに格納された情報を取り出すための専用のデコーダが必要であったことから部品・製品管理や在庫管理などの産業分野での利用が大半を占めていた．この数年で普及したスマーフォンに QR コードデコーダがアプリとして実装されたことから，ウェブサイトへのアクセス，入場券，決済サービス，アカウントの個人情報の伝達などの様々な用途での利用が広まっている．

QR コードは高い認識率を誇るが，単に黒と白のモジュールが並べられた画像データであるため，それに格納された内容を人は直接データとして解釈できない．つまり，利用者は QR コードの全体もしくは一部が書き換えられていたとしても気づくことが出来ない．また，QR コードに格納されたデータは正しいデータであると，むやみに信用してしまう利用者も多い．これを利用して，悪意のあるものが偽装した QR コードを作成し，それを読み取った利用者の不用意な操作により悪性サイトに誘導することが問題となっている．実際に，QR コードを決済に用いることが一般化した中国では，店舗側に表示された QR コードを第三者が張り替えることで，不正送金させる事件が発生している [2]．また，個人もしくは企業が QR コードを作成する際には，QR コードを作成するシステム開発会社への依頼，もしくは QR コード作成ソフトや QR コード作成サイトの利用が一般的である．それらの会社やサイトが悪意を持って正規の QR コードに偽装した QR コードを作成し，配布することも十分に考えられる．一方で，偽装された QR コードはそれを読み取ると必ず悪性サイトに誘導される．そのため，QR コードの設置者が事前にあるいは定期的に，QR コードを読み取って情報を確認すれば，偽装 QR コードの発見は容易であり早い段階での対策が可能である．また，利用者も読み取った情報を注意深くすることで，偽装を発見できる．決済サービスや認証サービスなどでの利用が進む QR コードの安全性を議論する上で，まだ明らかになっていない脆弱性や攻撃方法を発見することは非常に重要である．

本稿では，QR コードの構成方法，特に利用される誤り訂正符号の性質を利用して，発見が容易でない偽装 QR コー

[1] 神戸大学大学院工学研究科
[a)] okuma@stu.kobe-u.ac.jp

ドの作成方法を与える．具体的には，大きい確率で正規の情報を出力し，小さい確率で別の情報，すなわち悪意のある情報を出力する QR コードを作成する．悪意のある情報が出力される確率が小さいため，QR コードの設置者が数回確認した程度では正規の情報が出力される可能性が高く，QR コードが偽装されていることに気づくのは困難である．小さい確率で別の情報を出力させるために，QR コードの構成に用いられる誤り訂正符号の性質を利用する．誤り訂正符号はデータに生じた誤りを修復して正しいデータに復元するために用いられる技術であり，QR コードでは，QR コード画像に生じた汚れや撮影時の影などが原因で生じる誤りを訂正するために用いられる．誤り訂正符号を用いて訂正可能な誤りの量は決まっており，それを超える誤りが生じた場合には，別の情報が出力されること（誤訂正）が起こる場合がある．

著者らは，文献 [1] で，誤り訂正符号の性質を利用して，小さい確率で悪意のある情報を出力する偽装 QR コードの構成方法を提案している．前回の報告では，一つのモジュールにドット，あるいは輝度値を変えるなどして誤訂正の確率を制御し，一定の確率で悪性サイトに誘導する QR コードの作成例を与えた．しかし，一つのモジュールを変化させる（雑音を付加する）ことが偽装 QR コード作成における本質ではなく，特定のモジュールを誤認識するように QR コード全体を再構成することが，提案する偽装 QR コードの特徴である．モジュールへの雑音付加以外にも，QR コードを切り出すためのマーカの位置を工夫したり，特定のモジュール群を歪めることでも実現可能である．

本稿では，文献 [1] で提案した偽装 QR コードの作成方法を一般化するとともに，一部のモジュールだけでなく QR コード全体に変化を加えることで，通常の QR コードとの識別がより困難な偽装 QR コードの構成方法を提案する．さらに，我々が提案する偽装 QR コードだけでなく，様々な方法で偽装された QR コードへの対策について考察する．

2. RS 符号

RS 符号は誤り訂正符号の一つで，QR コードや CD，デジタル放送，衛星通信などに利用されている．誤り訂正符号とは，送信する情報に冗長な情報を付加することによって，通信中に発生した誤りを訂正し，正しいデータを復元する符号である．通信と同様に，情報の信頼性が重要である QR コードのような記録媒体にも利用されている．RS 符号は符号語を一定数のビットのかたまり（シンボル）の集まりで表し，各シンボル単位で誤り訂正を行うので連続して起こるビット誤り（バースト誤り）に強い．QR コードを読み取る際に生じる誤りは，QR コードの一部の汚れや損傷など連続して起こる誤りが多く，RS 符号と相性がよい．

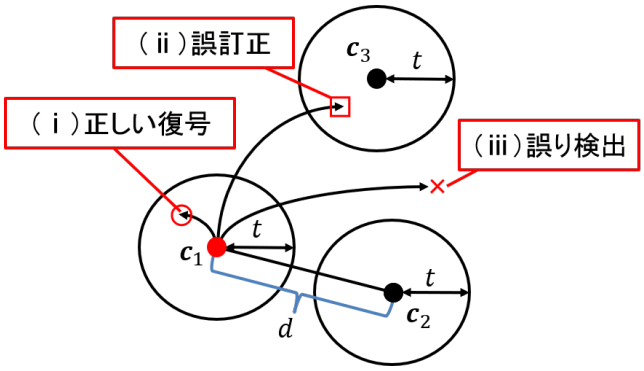

図 1　誤りの数と復号結果の関係

Fig. 1　The relation between number of errors and decoding result

符号長 n，情報点数 k の (n,k)RS 符号は，k シンボルで表された情報系列を $n(>k)$ シンボルの符号語系列に符号化する．QR コードでは，8 ビットを一つのシンボルとして扱い情報系列を符号化する．QR コードデコーダを用いて読み取った系列から元のデータを推測する（復号という）際には，ユークリッド復号法 [3] が用いられている．ユークリッド復号法は代数計算のみで誤りの訂正が可能であり，計算量が小さく，QR コードの高速読み取りを実現に貢献している．(n,k)RS 符号において，ユークリッド復号法を用いて訂正可能な誤りの数は設計距離 $d=n-k+1$ により保証されており，訂正可能な誤りシンボルの数（誤り訂正能力）t は，

$$t=\left\lfloor \frac{d-1}{2} \right\rfloor = \left\lfloor \frac{n-k}{2} \right\rfloor \qquad (1)$$

で与えられる．(n,k)RS 符号を用いると，QR コードの読み取り時にノイズが発生しても，t シンボル以下の誤りであれば元のデータを復元することができる（図 1(i)）．また，t シンボルを越える誤りが生じ，受信したデータが元のデータ以外の任意の符号語から距離が t 以下となる場合はその符号語に復号される（誤訂正）（図 1(ii)）．それ以外の場合は誤り検出となる（図 1(iii)）．RS 符号の性質上，誤り訂正能力を超える誤りが発生したとしても，誤訂正が起こる確率は非常に小さく，ほとんどの場合は誤り検出となる [3]．

3. QR コード

3.1 QR コードの性質

QR コードは格納されている URL 等の情報を高速かつ正確に読み取りが可能な二次元コードで，1994 年に株式会社デンソーの開発部門（現在は分離し株式会社デンソーウェーブ）によって開発された．QR コードに格納する情報は RS 符号により符号化され，RS 符号のパラメータを変更することで，誤り訂正レベルを L，M，Q，H の四段階に設定可能である．最も高い誤り訂正レベル H では QR コードの約 30 %が損傷していても元のデータを復元する

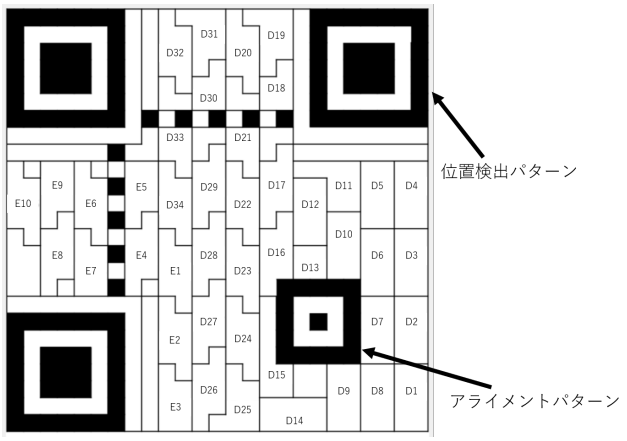

図 2 QR コードの構成
Fig. 2 Structure of QR code

ことができる．誤り訂正レベルを上げれば誤り訂正能力は向上するが，誤り訂正に必要な冗長部分が増えるので，格納できる情報の容量が小さくなる．一般的に使用されているモデル 2 の QR コードでは 1 型（21 × 21 モジュール）から 40 型（177 × 177 モジュール）までの型番が用意されており，バイナリデータは 2953 バイトまで格納することができる [4]．

3.2 QR コードの構成

二次元コードである QR コードはデータを白黒の正方形のモジュールで表し，それ配置する方法で作成される．QR コードは位置検出パターン，アライメントパターンなどの機能パターンと，情報ブロック，誤り訂正ブロック，形式情報などの符号化領域で構成される [5]．情報ブロックは，格納する情報コードと埋め草コードからなる．埋め草コードは情報コードが QR コードの容量に満たないときに付与される無為なデータである．図 2 は，QR コードの構成である．

QR コードは次の手順で構成される．

Step 1: QR コードに格納する文字列を指定の文字コードを用いて二進数に変換する．

Step 2: 二進数化したデータの先頭にモード指示子，文字数指示子を，末尾に終端パターンを付加し，情報コードとする．

Step 3: 情報コードのシンボル数が格納する QR コードの情報ブロックの容量に満たなければ，足りない分だけ埋め草コードを付加する．

Step 4: 情報コードと埋め草コードを合わせた情報ブロックを RS 符号で符号化する．

Step 5: 作成した符号語を型番ごとの仕様に従って QR コードに配置し，マスク処理を施す．

Step 3 で付加される埋め草コードは，無為なデータであり，その大きさが等しければであれば任意に変更できる．Step 5 で施すマスク処理はモジュールの明暗の偏りを無くすためであり，必要がなかればマスク処理をしないことも可能である．

3.3 QR コードのデコード

QR コードデコーダの多くは Google が開発・提供している QR コードライブラリ “ZXing”[6] を利用している．ZXing を用いた QR コードのデコードは下記のように行われる．

Step 1: 撮影した画像ファイルをグレースケールに変換する．画像を分割し，部分ごとに輝度値の閾値を決定し，明暗の二値化を行う．

Step 2: モジュールの明暗の比率から位置検出パターンとアライメントパターンを検出する．

Step 3: 位置検出パターンを利用して QR コードのサイズやモジュールのサイズを取得し，モジュールの中心座標を計算する．また，アライメントパターンよりモジュールの中心座標計算のズレを補正する．

Step 4: 各モジュールの中心座標の明暗を元に，そのモジュールの明暗を判別する．

Step 5: 取得したビット値においてマスク処理を解除して受信語を構成し，誤り訂正を通してデータを復元する．

汚れや影，光の反射などの影響によって，Step 3 で決定したモジュールの中心部分の明暗の読み取り結果が誤る場合がある．QR コードは一部の誤りを訂正する能力が備わっているので，訂正能力を超えた誤りでなければ，Step 5 で格納された情報を正確に復元することができる．

4. 偽装 QR コードの符号語の構成

本章では，文献 [1] で提案した偽装 QR コードの構成方法を一般化し，確率 p で格納した情報 A を出力し，確率 $1-p$ で異なる情報 B を出力する QR コードの構成方法を説明する．異なる二つの情報を出力するためには，片方が誤訂正が生じたときに出力されるように設計すればよい．

4.1 誤訂正の生起条件と格納する系列の構成

RS 符号では，任意の二つの符号語間距離の最小値は d で保証されているが，特定の二つの符号語の距離はそれよりも大きくなる場合がある．誤訂正が生じるためには，情報 A の符号語に誤りが生じて，情報 B の符号語との距離が近づく必要がある．二つの情報 A，情報 B の符号語間の距離が最小距離 d だとしても，都合よく情報 B の符号語に近づく誤りが生起することはほぼない．すなわち，情報 A が格納された QR コードを読み取って，誤訂正により情報 B を出力する確率は非常に小さい．したがって，一般的な QR コードの構成方法では格納した情報と異なる情報を出力させることは困難である．そこで，予め訂正能力の限界のシンボル数まで，符号語 $\boldsymbol{c}_A$ のシンボルを符号語 $\boldsymbol{c}_B$ のシンボルに置き換えた系列を QR コードに格納することで，

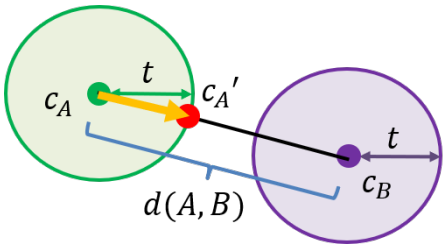

図 3　符号語の関係
Fig. 3　The relation between codewords

誤訂正を起こしやすくすることを考える．

誤訂正を用いて偽装 QR コードを構成する手順は以下の通りである．

Step 1: 情報 A と情報 B の符号語 $\boldsymbol{c}_A$ と $\boldsymbol{c}_B$ を RS 符号化より生成する．符号語 $\boldsymbol{c}_A$ と $\boldsymbol{c}_B$ の距離を $d(A,B)$ とする．

Step 2: 符号語 $\boldsymbol{c}_A$，$\boldsymbol{c}_B$ 間で異なる $d(A,B)$ シンボルの中から符号語 $\boldsymbol{c}_A$ の任意の t シンボルを選択し，符号語 $\boldsymbol{c}_B$ の同じ位置のシンボルに置き換えて，系列 $\boldsymbol{c}_A^{'}$ を作る．

Step 3: QR コードの構成方法に従って系列 $\boldsymbol{c}_A^{'}$ を白黒のモジュールとして格納する．

図 3 は，偽装 QR コード作成時に考える各系列の関係を図示したものである．QR コードを読み取った際に，Step 2 で選択しなかった $d(A,B)-t$ シンボルに，誤りが生じて符号語 $\boldsymbol{c}_B$ の同じ位置のシンボルと同値となる場合に誤訂正が起こり，符号語 $\boldsymbol{c}_B$ が復号結果として出力される．

4.2　誤訂正の生起確率の制御

前節の方法より，t 個のシンボルを予め情報 B の符号語と一致させた．誤訂正が生じるには他に $d(A,B)-t$ シンボルの誤りが必要で，それらのシンボルが別の符号語と一致しなければならない．しかしながら，QR コードの読み取りにおいて，モジュールの色を間違って読み取ることは起こりにくいため，二つの符号語間で $d(A,B)-t$ 個のシンボルが一致する確率は低い．

そこで，著者らは文献 [1] で，作成した偽装 QR コードに対してノイズを付加して誤訂正の生起確率を制御する方法を与えた．例えば，白モジュールから黒モジュールへの誤りについて考えると，白モジュールに対して小さな汚れを付加する，またモジュールの輝度値を変えるなどのノイズを意図的に付与することで，読み取りの際に誤りが生じやすくなる．これを利用して，誤訂正の生起確率を任意に調節することが可能である．しかし，これらのノイズは注意深く観察することで発見できるため，偽装されていると判断されやすい．次章では，注意深く観察したとしても，通常の QR コードと判別することが困難な QR コードの作成方法を提案する．

図 4　ドット状の汚れで誤りを制御する偽装 QR コード
Fig. 4　A fake QR code with dot-type noize

4.3　偽装 QR コードの作成例

文献 [1] で作成した偽装 QR コードの例を紹介する．図 4 は，高確率で URL1（http://www.netbk.co.jp）を出力し，低い確率で異なる URL2（http://www.netbktco.jp）を出力する偽装 QR コードである．これは，特定のモジュールにドット状の汚れを付加し，その汚れの輝度値を変化させている．(44,28) RS 符号を利用する 2-M 型 QR コードを用いた．(44,28) RS 符号の最小距離は $d=17$ であり，誤り訂正能力は $t=8$ である．

Step 1: URL1 と URL2 をそれぞれ情報コードに変換した後，RS 符号により符号化し，符号語 $\boldsymbol{c}_A$ と符号語 $\boldsymbol{c}_B$ を生成する．

Step 2: 符号語 $\boldsymbol{c}_A$，$\boldsymbol{c}_B$ 間で異なる 17 シンボルの中から，符号語 $\boldsymbol{c}_A$ の任意の $t=8$ シンボルを選択し，符号語 $\boldsymbol{c}_B$ の同じ位置にあるシンボルに置き換えて，系列 $\boldsymbol{c}_A^{'}$ を作成する．

Step 3: QR コードの構成方法に従って系列 $\boldsymbol{c}_A^{'}$ を白黒のモジュールとして格納する．

Step 4: 系列 $\boldsymbol{c}_A^{'}$ と符号語 $\boldsymbol{c}_B$ 間で異なる $d-t=9$ の中から $d-2t=1$ シンボルを選択し，符号語 $\boldsymbol{c}_B$ のシンボルと一致するようにノイズ（例：ドット状の汚れ）を付加する．

ドット状の汚れを付加する場合，カメラで認識可能な大きさの汚れを付加する必要があるため，注意して観察すれば QR コードが偽装されていると視認可能である．

5.　偽装の判別が困難な QR コード

先に，文献 [1] では，モジュールの一つにドットを付加する，あるいはそのモジュールのみ輝度値を変える等，特定のモジュールに雑音を付加することにより，偽装 QR コードを具体的に実現した．しかしながら，一つのモジュールへの雑音付加が偽装 QR コード作成の本質ではなく，特定のモジュールを誤認識するように QR コード全体を再構成

することが，提案する偽装 QR コードの特徴である．QR コードのモジュールを正確に切り出すための各マーカの位置を工夫したり，特定のモジュール群を若干歪めることで実現可能である．本章では，一つのモジュールに雑音を加えるのではなく，偽装 QR コードの発見（通常の QR コードとの識別）を困難にするための一つの方法を与える．

QR コードの応用として，デザイン QR コードが普及している．デザイン QR コードとは，一般的な白黒のみのコードではなくカラフルで画像や文字が入っている QR コードのことで，Web サイト上のデザイン QR コード作成ソフトによって誰でも簡単に作成することができる [7]．本節では，画像と偽装 QR コードを重ね合わせ，高確率で URL3（http://www.kobe-u.ac.jp）を出力するが，背景となる画像の影響によって異なる URL4（http://www.nara-u.ac.jp）を出力する偽装 QR コードの構成を示す．偽装 QR コードの作成には，文字数の関係で 2-L 型 QR コードを用いた．(44,34) RS 符号を利用し，最小距離は $d = 11$ であり，誤り訂正能力は $t = 5$ である．

Step 1: URL3（http://www.kobe-u.ac.jp）と URL4（http://www.nara-u.ac.jp）の二つの情報を符号化し，符号語 $\boldsymbol{c}_A$，$\boldsymbol{c}_B$ をそれぞれ生成する．

Step 2: 符号語 $\boldsymbol{c}_A$，$\boldsymbol{c}_B$ 間で異なる $d(A,B) = 12$ シンボルの中から，符号語 $\boldsymbol{c}_A$ の任意の $t = 5$ シンボルを選択し，符号語 $\boldsymbol{c}_B$ の同じ位置にあるシンボルに置き換えて，系列 $\boldsymbol{c}_A^{'}$ を作成する．

Step 3: QR コードの構成方法に従って系列 $\boldsymbol{c}_A^{'}$ を白黒のモジュールとして格納する．

Step 4: 作成した QR コードと同じ大きさの画像を基に，QR コード上の黒モジュールに対応する座標では画像上の輝度値を小さく（暗く），また，白モジュールに対応する座標では画像上の輝度値を大きく（明るく）し，構成する．

Step 5: 系列 $\boldsymbol{c}_A^{'}$ と符号語 $\boldsymbol{c}_B$ 間で異なる $d(A,B) - t = 7$ の中から $d(A,B) - 2t = 2$ シンボルを選択し，符号語 $\boldsymbol{c}_B$ のシンボルと一致するように背景画像の輝度値を制御する．

Step 5 での輝度値の制御により，出力される情報の確率を制御することができる．

本稿で作成した偽装の判別が困難な QR コードを図 5 に示す．また，URL3（http://www.kobe-u.ac.jp）のみを出力する偽装されていない QR コードを画像と重ねたものを図 6 に示す．

図 5 偽装された QR コード
Fig. 5 An example of fake QR code based on design QR code

図 6 偽装されていない QR コード
Fig. 6 An example of design QR code

6. 偽装 QR コードへの対策

携帯電話の普及以降，さまざまな QR コード読み取りアプリケーションが開発・公開されているが，セキュリティ面で対策がなされているアプリケーションは少ない．アプリケーション側で対策されるまでは，利用者が QR コードが安全であるかを見極めなければならない．利用者が取ることができる対策の一つは出力された URL 等の情報を注意深く確認することである．QR コードの読み取りアプリケーションの中には，取得した URL を表示せずに直接リンク先のホームページに遷移するものがあるそのような設定で利用する際は QR コードに格納された情報を確認することができないため非常に危険である．読み取った情報が QR コードに併記されている URL と完全に一致しているか注意深く確認することが重要である．

アプリケーション側で取られる対策もいくつか考えられる．単純なものでは，見た目が怪しい QR コードの読み取りを行わないという機能を付けることが挙げられる．機械学習を用いてアプリケーションに QR コードの形状を学習させ，通常の QR コードと判定しなかった場合は読み取りを中止する，または利用者に警告する機能をつけるという方法が考えられる．しかしこの対策では，デザイン性を重視した QR コードなど多様な外見を持つ QR コードの読み取りを制限する可能性がある．また，意図せず汚れが付着した QR コードも読みとらなくなる可能性があり，QR

コードが本来持つ汚れていても情報を復元できるという機能が意味をなさなくなる．他の対策としては，読み取ったURLからリンク先のWebページが安全であるかどうかをチェックするという機能の追加が挙げられる．QRコードに格納されている情報が短縮URLであった場合，利用者はそのURLを見てもリンク先が安全であるか危険であるか判断することができない．この対策では，たとえ短縮URLであってもリンク先を第三者がチェックするなどして安全性を確認することができる．以上の対策を講じるなどして，様々なQRコードの偽装をアプリケーション側で防いでいく必要がある．

7. まとめ

文献[1]では，一つのモジュールに雑音を付加することで，偽装QRコードを具体的に実現した．しかしながら，一つのモジュールへの雑音付加が偽装QRコード作成の本質ではない．本稿では，背景画像とQRコードを組み合わせたデザインQRコードを元に，特定のモジュールを誤認識するようにQRコードを再構成し，偽装されていることの識別が困難なQRコードの具体的な作成方法を与えた．提案手法は，ほとんどの場合，正規のサイトに誘導されるが，小さな確率で悪性サイトに誘導される偽装QRコードの作成を可能とする．それゆえ，再現性が難しく，発見を遅らせることになり，被害の拡大となることが予想される．また，偽装されたQRコードと通常のQRコードを視認によって区別することは困難であり，偽装されたQRコードの発見は必ずしも容易ではない．特に悪意のあるQRコード作成者によって，このような偽装されたQRコードが作成，配布された場合，大きな被害が想定される．今後はQRコードの信頼性について問題とするとともに，各人はより注意深く利用する必要があると言える．

参考文献

[1] 大熊浩也, 瀧田愼, 森井昌克, “悪性サイトに誘導するQRコードの存在とそれを利用した偽造攻撃”, 信学技報, vol. 118, no. 109, ICSS2018-6, pp. 33-38, 2018年6月

[2] 牧野武文，“シールを貼るだけのお手軽詐欺 アリペイの偽QRコードで1万元を盗む”，https://the01.jp/p0005594/，sprout，August 21 2017 (参照 2018-08-20).

[3] 今井秀樹, “符号理論”, 電子情報通信学会, pp.155-173, 1990.

[4] 日本工業規格, JIS, X0510, 二次元コードシンボル—QRコード—基本仕様, 2004.

[5] 池田和興, “例題が語る符号理論　BCH符号・RS符号・QRコード”, 共立出版, 2007.

[6] “ZXing (“Zebra Crossing”) barcode scanning library for Java, Android”, https://github.com/zxing/zxing (参照 2018-08-20).

[7] Unitag, https://www.unitag.io/qrcode (参照 2018-08-20).