

Boston University Electrical and Computer Engineering EC521 -- Cyber Security Project -- Final Report

QR Code Security



Team Members:

George Kent-Scheller, georgeks@bu.edu
Sam Krasnoff, krasnoff@bu.edu
John Kircher, kircherj@bu.edu
Julian Padgett, jpadgett@bu.edu

For access to the Github repository, please contact: georgetks@gmail.com

Description

QR codes are a machine-readable code consisting of an array of black and white squares, typically used for storing URLs or other information for reading by the camera on a smartphone. Since the beginning of the COVID-19 Pandemic, usage of QR codes has increased to eliminate the need for physical human interaction and reduce the risk of infection. Today, QR codes have become ubiquitous in many facets of modern life. However, QR codes are not very safe or secure to use. Phishing attacks can be executed by getting victims to scan QR codes in the same way that an attacker could get a victim to click on a dangerous link. While many people are educated on the importance of not clicking on suspicious links, there seems to be less public awareness of the importance of not scanning suspicious QR codes.

Approach / Measurement

In order to understand the extent of the problem, our team created several fake flyers with our own QR codes, and posted over 80 of them around the BU campus and surrounding areas. These flyers included a sign-up form for a nonexistent basketball tournament, a fake Amazon gift card giveaway, and free tickets to a fictional nightclub event.







Figure 1: This image depicts our Amazon, Party, and Basketball QR code flyers

Over one week, 55 people scanned the Amazon flyers, 10 people scanned the basketball flyers, and 5 people scanned the club flyers. After scanning the QR codes attached to the flyers, multiple steps were taken to redirect users and record their data:

- 1. After scanning a QR code, a javascript program accessed data from the user including:
 - a. Google Analytics Cookies, IP, QR Code Scanned, Device, Operating System, Search Engine, and Time.
- 2. Users are then redirected to a website hosted using Heroku. This website included a brief description of our project, as well as the ability to fill out a survey about how they accessed the QR code, their awareness and understanding of cyber security, and the ability to leave any notes or thoughts about the experiment. (A view of our website can be seen in Figure 9, page 8)

- 3. POST requests are made to another javascript program (app.js) with all of the collected data.
- 4. This data is then added as a document to a collection in a FireStore database. Furthermore, if users filled out the survey via the web page, their data was also added to another collection in FireStore.

We also decided to add a control to our data collection. This control was a fake game-link that was shared via Facebook student groups, and Reddit student groups.



Figure 2: This image depicts our President Brown The Game advertisement

The link was supposed to act as a control to see how many more people were susceptible to QR codes over random "sketchy" links. However, as seen from our data in the next section, more people clicked on the link than scanned QR codes in total. Ultimately, we also included the data in our research because it still exposes the vulnerability that people will scan/click on images and codes with no thought to security risks.

Results

In total, we received 184 documents through our FireStore database. These included 70 (38%) from various QR codes, and 114 (62%) from the game link. 78.5% of the total scans were from Amazon, 14.3% from the basketball flyers, and the final 7.2% from the party fliers. Histograms showing the distribution of data are shown below in

Figure 3. This data allows us to draw numerous conclusions. Among the flyers, individuals were much more likely to scan something with immediate monetary gain, rather than arbitrary events that may just be interesting to the individual. We were genuinely surprised, as the Amazon flyer was by far the least official looking. In terms of carrier/IP breakdown, 44% of the data came from users directly on the BU network. The university has sent out numerous emails warning users to be wary of suspicious links, and other connections to digital mediums, like QR codes, which were clearly ignored. The rest came from various LTE carriers spread evenly over most of the major US companies. The most traction we got was from the game link, posted in a private Facebook group and on BU reddit. Though the post did seem genuine, social media accounts are often hacked and exploited, and users should always be careful of clicking on unknown links, as there was a clearly fake hyperlink. Overall, this experiment shows a consistent lack of good digital security practices, and that to protect the lowest knowledge user, there must be a secure alternative.

Furthermore, we also write a python script accessing the xlwt library to get all of our data from FireStore and write it to an Excel sheet for a more simple analysis. An example/screenshot of this sheet can be seen in figure 4.

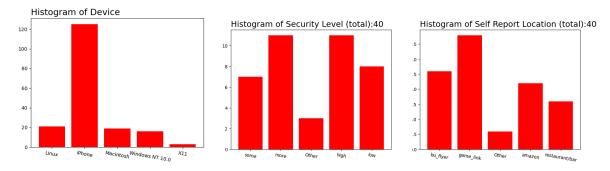


Figure 3: These graphs represent data collected from individuals who scanned the QR codes and visited our website

Figure 4: This is our Excel Sheet holding all of our data parsed from JSON to a list in python

While we did receive favorable results from our experiment, we could have had many more scans had we placed QR codes in more vulnerable locations. For example, if we replaced restaurant menu QR codes with our own, or replaced official BU QR codes with our own. However, we did not want to intrude or adversely affect businesses or do anything illegal. Finally, we also received various forms of feedback through our survey that was included on the website. Though we had many responses (40 total, as seen from histograms in Figure 3), only 15 included their thoughts which can be seen below in Figure 5. These ranged from helpful comments to some of our peers expressing their surprise and frustration with the experiment.

Other: "rip my tuition"

Game link: "Did someone hack me"

Bu flyer: "nice"

Amazon: "Really interesting, when I saw this I pop up I was like I've never rlly thought about that"

Amazon: "False advertising bullshit"

Game_link: "This is kind of a stupid experiment. Why would I expect a phasing scam on a student Facebook group with user authentication required to enter?"

restaurant/bar: "Okay Yes"

Bu_flyer: "I have client sided protection that would halt and alert before anything was executed or downloaded."

Amazon: "Hmm"

Game_link: "Interesting"

Game_link: "I love it!"

Bu_flyer: "Cool. Good way to see how dumb people are (clearly I am lol)"

Amazon: "Convincing"

restaurant/bar: "Yes"

Bu_flyer: "this exp gave me a descent scare. its a good way of spreading important information"

Figure 5: This is a table showing 15 written responses received from our survey

Mitigation

Our primary mitigation was the website. Although campus IT routinely sends out emails warning people about specific security risks, our project proves that people still don't understand the dangers of blindly clicking or scanning. This website not only helped us see what people were most likely to scan but also was meant as a prototype for how an I.T. website could work. If there are more fake websites that exist to demonstrate security vulnerabilities without actually exploiting them, we believe that it would improve security awareness. Many of our survey comments reflected peoples appreciation of the learning experience.

The secondary purpose of our project was to make QR codes more secure. Our solution is to generalize https. We believed that this security standard would help preserve the power of QR codes. One of the really cool things about QR codes is their ability to store raw bytes. This meant that we wanted to write a program to sign the actual QR codes rather than just checking for https after the scan. However there is a big issue with our mitigation.

Although it is not well documented, there is a lack of standards in QR code scanners. Many lower quality QR code scanner libraries are not able to scan the largest types of QR codes. This is an issue because the ability to write programs, music files, or images to a QR code cannot take full advantage of the 2953 bytes offered by the version 40 QR codes. This makes the space on the QR code at premium. In our example, we minimize space used by utilizing SHA1 hashing for our signing. The user could increase security by writing a more robust QR code scanning library that supports the largest sized QR codes. They could also implement a system where they have a signature on one QR code and use the other for data only.

For our version we used RSA signing on a SHA1 hash of the data. We then appended the signature to the end of the data that we wanted to put on the QR code. The data was then turned into a QR code and scanned by the user. The QR code was separated into data and signature sections. The signature was then verified against the

data with a public key. There would be a certificate authority who would be able to sign important QR codes. Additionally there would be the option for one party to provide non-repudiability of data on a QR code to a separate party. Figure 7 (below) shows prototypes of the secure QR code, with an embedded hash, but no authority to reference. In a proper implementation, the user would scan, confirm trust, and then forward the user to their data. If it was an untrusted source, an intermediate page would interrupt the process, similar to an HTTP (non-secure) link on a desktop browser.

As the images show below, we were able to make a decently robust proof of concept QR code generator. In Python, we used a QR code package downloaded through PIP, as well as the built in crypto library. Running the main generator function will take the inputted URL/information and append a SHA1 hash of a trusted source.

```
| George | A generals | Quencials | Quenci
```

Figure 6: This is our generate py file that generates the keys and encodes the url and binary



```
<class 'Crypto.PublicKey.RSA.RsaKey'> Public RSA key at 0x7F52823F86D0
https://theoldpurple.com True
[Finished in 6.212s]
```

Figure 7: This is our signed QR code output for the URL "https://theoldpurple.com" next to the scanner method with the first return value the data and the second the verification status.

```
sign(message, priv_key, hashAlg="SHA-256"):
 from Crypto.PublicKey import RSA
                                                                               global hash
 from Crypto.Cipher import PKCS1_OAEP
 from Crypto.Signature import PKCS1_v1_5
                                                                               signer = PKCS1_v1_5.new(priv_key)
from Crypto. Hash import SHA512, SHA384, SHA256, SHA, MD5
                                                                               if (hash == "SHA-512"):
from <u>Crypto</u> import Random
                                                                               elif (hash == "SHA-384"):
/ def newkeys(keysize):
                                                                               elif (hash == "SHA-1"):
    random_generator = Random.new().read
    key = RSA.generate(keysize, random_generator)
    private, public = key, key.publickey()
    return public, private
                                                                               return signer.sign(digest)
                                                                           def verify(message, signature, pub_key):
return RSA.importKey(externKey)
                                                                               signer = PKCS1_v1_5.new(pub_key)
def getpublickey(priv_key):
def encrypt(message, pub_key):
    cipher = PKCS1_OAEP.new(pub_key)
    return cipher.encrypt(message.encode('utf-8'))
def decrypt(ciphertext, priv_key):
                                                                               return signer.verify(digest, signature)
     cipher = PKCS1_OAEP.new(priv_key)
    return cipher.decrypt(ciphertext)
```

Figure 8: This is our RSA.py that uses RSA encrypt and decrypt to verify and sign the keys

Our Website:



Figure 9: This is an image of our website hosted using Heroku