A Review of Detection Methodologies for Quick Response code Phishing Attacks

Sikiru Subairu
Department of Cyber Security Science
Federal University of Technology
Minna, Nigeria
islam4life@futminna.edu.ng

John Alhassan
Department of Cyber Security Science
Federal University of Technology
Minna, Nigeria
jkalhassan@futminna.edu.ng

Shafii Abdulhamid
Department of Cyber Security Science
Federal University of Technology
Minna, Nigeria
shafii.abdulhamid@futminna.edu.ng

Joseph Ojeniyi
Department of Cyber Security Science
Federal University of Technology
Minna, Nigeria
ojeniyija@futminna.edu.ng

***Abstract*-Recently, phishing attacks have taking a new dimension with the addition of quick response code to phishing attacks vectors. Quick response code phishing attack is when an attacker lures its victims to voluntarily divulge personal information such as password, personal identification number, username and other information such as online banking details through the use of quick response code. This attack is on the rise as more and more people have adopted mobile phone usage not just for communication only but to perform transaction seamlessly. The ease of creation and use of quick response code has made it easily acceptable to both provider of goods and services and consumers. This attack is semantic as it exploits human vulnerabilities; as users can hardly know what is hidden in the quick response code before usage. This study reviewed various methodologies that earlier researcher have used to detect this semantic-based attack of phishing. The strength of each methodology, its weakness and general research gaps identified.**

***Keywords*- phishing attack, password, quick Response Code, username**

I. INTRODUCTION

A quick Response (QR) code which is two dimensional based matrix barcodes has now become a popular phishing and malware attacks vector [1]. This code is easy to produce and deploy, hence it has recently become part of daily lives usages ranging from billboard advertisement to access potential buyers of goods and services and uniform resource locator (URL) encoding in making availability of information for navigation.

Figure1: QR-Code Sample

These merits of QR code is now being used as an attack vector [2] by phishers to fraudulently direct users to phishing sites to obtain vital financial personal data from them or have malware downloaded to their devices [3]. QR codes ability to facilitate easy sharing of varieties of data such URL, products details, phone number, business cards, email addresses and other information, coupled with ease of creation and usage has made its uses on the rise [4].

This technology application now cut across so many human endeavors such as mode of payment, authentication, access control, web navigation, ticketing, advertisement and host of other usages [5]. Researchers have opined that QR code is an excellent medium to launch phishing attacks and malware distribution as the content in the code is opaque and can only be read by machine, thus a user will find it too difficult to be able to differentiate between a benign QR code and malicious QR code. The users sometimes build their trust on the fact that trusted brand organization has QR code placement on their advertisement [3]. Numerous solutions to mitigate QR code phishing attacks have been proposed; however, phishing as threat is increasing and becoming a common fraud to commit e-crime [6].

II. QR CODE PHISHING

Financial fraud perpetrated through QR code has been on the rise, recently about US$13 million was reported stolen through QR code scam, and 900,000 Yuan was also reported lost to QR code scam in China [7]. The original QR code by the merchant was simply

978-1-7281-5467-1/20/$31.00©2020IEEE.

replaced with malicious QR code which then steals personal data upon scanning [7]. The simplicity of creating QR code and its deployment, coupled with its content being hidden from what human eye can verify as being malicious or benign; has made cybercriminals adopt QR code as phishing attack vector [2].

Ubiquitous nature of QR code has opened serious security challenges such as phishing and malware attacks which can launch simply by concealing malicious URL in QR codes [5]. Malicious QR code was first discovered in September 2011 by kaspersky laboratory, in which when a user scan such code, the user will be directed to a malicious website whereby malicious files will be downloaded to such device unknowing to the user [8].

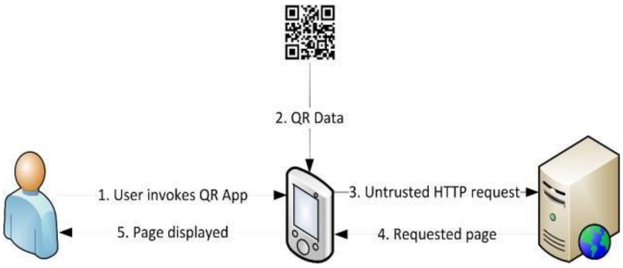

Figure 2: Webpage Request Using QR Code [9]

Obfuscating URL destination in QR code and embedding this code in spam email, has proven to be an effective attack vector for phishing and malware, as it easily evade malicious link detection algorithm as being applied to email traditionally [9]

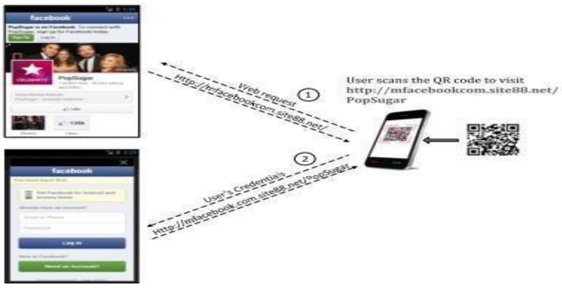
facebook
PopSugar
facebook
Log in
Web request
Http://mfacebookcom.site88.net/
1
User scans the QR code to visit http://mfacebookcom.site88.net/PopSugar
2
User's Credentials
Http://mfacebook.com.site88.net/PopSugar

Figure 3: QR Code phishing Attack [2]

III. QR CODE PHISHING DETECTION METHODOLOGIES

Detecting QR code Phishing attacks has not been widely discussed in available literatures unlike phishing generally. Previous works focus more on the acceptability of this technology and dangers inherent in it such as it being a vector of phishing attack and malware exploit. Few works that focus on QR code phishing detection are as follows with their contributions and limitations.

[1] proposed a solution to QR code phishing detection by leveraging on two API which are; Google Safe browsing and PhishTank. Google Safe Browsing API enables application to check status of URL against updated list of phishing and malware websites. PhishTank is a database that contains blacklists of phishing URL made up of manually verified and checked websites. However API can be manipulated [10], so it cannot be relied upon, likewise this solution cannot handle zero day attack of phishing, as phishing URL at times do not last for long. Research neither presents the result of the proposed solution in terms of detection rate of QR code phishing nor benchmarks against other existing work. Rather a comparative result of user experience using this solution and other work in terms of user warnings was presented.

[2] proposed a methodology of URL Matcher, where a QR code is extracted using an extractor from a crawled websites and then compared with the domain feeds from URL blacklist. Research failed to realize that QR code with malicious intent may not necessarily be online for it to be crawled as phishing sites does not last long [11]. Effectiveness of this solution cannot be verified in terms of detection rate as basis for comparison were not established, no QR code phishing dataset formulated in the research.

[12] proposed a detection method for malicious QR code using tamper detection system with a digital signature and wet papercode approach. QR code contents hash function is evaluated, public and secret key formulated and QR code content is encrypted with secret key. Wet paper code is utilize in embedding encrypted message into QR code and forwarded to receiver. Receiver decrypt using public key, evaluate the hash function and compare hash values to determine if the message has been tamper. Research implementation is not carried out; hence there is no evaluation of this proposed solution. Also the work focuses mainly on QR code generation without considering the usability of QR code technology.

[13] add to the security of QR code in terms of generation, authentication of QR code to be trusted, verification of online content of the code and finally isolate malicious content of QR code. Digital Signature Algorithm (DSA) asymmetric key cryptography plus SHA-1 was used for code generation, authentication and data integrity.

There is no evaluation of research in terms of detection rate and benchmarking proposed solution against existing work, rather the only evaluation done was in terms of time taken for the proposed system to generate QR code with secured system and decodes. This proposed solution cannot handle QR code phishing zero day attacks, and its usability may not be encouraging to user due to complexity of time and space. There is no publicly available QR code phishing dataset formulated in the research.

The research [3] proposed a design guideline for a secure QR code such as enhancing the visual QR code to mitigate modification of the code. Secondly, they also proposed embedding digital signature into QR code for user to be able to verify code source. No implementation carried out, and work focuses only on QR code generation and not detection.

[14] adopted machine learning approaches to detect phishing in QR Code. Methodology used in this work is machine learning with Naive Bayes algorithm for classification. Features selection for dataset was manually done with twenty instances base on host and lexical structure of URL phishing. Detection rate achieved in the work is 93.34%, though performance is encouraging but can be enhance with better methodology such as CNN, which learn features automatically from dataset with more instances of phishing attacks than manually features selection. No publicly available QR code phishing dataset formulation from the research.

[5] extensive comparative analysis of malicious QR code detection tools was carried out. Based on this work finding, it clearly established that more work is needed in the area of phishing detection in malicious QR code as the best detection rate is 88.33%. The Phishing URL in this work was obtained from the best three phishing databases publicly available; which are phishtank, openphish and malware domain; hence, result could be relied upon. However, no publicly available QR code dataset was formulated; research focuses not on detection model but rather on detection tools.

IV. REVIEW SUMMARY

Table 1.0 shows the different research works carried out on QR-code phishing. It pointed out the methodology used, results obtained and their limitations.

Table I. RELATED LITERATURE REVIEW

S/N	Reference	Problem Solved	Methodology	Result	Limitation
1	Yao & Shin, 2013	QR code phishing detection	Leveraging on two API which are;Google Safe browsing and PhishTank.	Solution provided by research is better in terms of user usability	API can be manipulated, no detection rate, TP, TN, FP, FN mentioned. No QR code phishing dataset formulated
2	Kharraz et al., 2014	Detecting QR code from crawled websites	Blacklisting: crawler extract QR code from crawled websites and then compare it with blacklisting database.		Malicious QR code may not necessary be online. A base for evaluating solution is not established. It cannot handle Zero day QR code phishing attacks. No QR code phishing dataset formulated
3	Ishihara &Niimi, 2014	Proposed a solution to detect changes in QR code contents.	Temper detection system with digital signature and wet paper code.	Work not implemented	Work focuses only on QR code generation and not on detection, using encryption without considering usability.

S/N	Reference	Problem Solved	Methodology	Result	Limitation
4	Bani-Hani, Wahsheh, & Al-Sarhan, 2014)	Secured QR code generation	Digital Signature Algorithm (DSA) asymmetric key cryptography plus SHA-1 was used for code generation, authentication and data integrity	Time taken for their system to generate QR code with secured system and decodes was encouraging	Research proposed solution cannot handle QR code phishing zero day attacks, and its usability may not be encouraging to user due to complexity of time and space.
5	Krombholz *et al*., 2015	Proposed a design guideline and digital embedment for secure QR code generation.	Embedding digital signature and enhancing QR code visual content.	Verification of QR code source and integrity.	Work focuses not on QR code phishing detection.
6	Alnajjar et al., 2016	Detection of phishing in QR Code.	Machine learning with Naive Bayes algorithm for classification. Features selection was manually done with twenty instances base on host and lexical structure of URL phishing.	Detection rate achieved in the work is 93.34%,	Performance is encouraging but can be enhanced with better methodology such as CNN which learn features automatically from dataset with more instances of phishing attacks than manual features selection. QR code phishing dataset not available
7	Dudheria, 2017)	Comparative analysis of QR code phishing tools.	Scanning malicious QR code URL obtained from phishtank, openphish and malware domain.	Best performed tool has 88.33%. Authors established need for more research in the area of detecting phishing in QR code.	Low detection rate and tools cannot mitigate zero day QR code phishing attacks.

V. FINDINGS FROM LITERATURE REVIEW

From the literatures review, several countermeasures have been proposed to solve this problem of QR code phishing attacks which arises from concealing malicious URL, though little were implemented [1, 13, 14, 5]. Prominent amongst them is the use of an algorithm which integrates both goggle safe browsing API and Phishtank API in detecting QR code phishing attacks, but the drawback in this solution is that, the users still have to scan the QR code before detection is initialized; hence user is still exposed to security threats. [3] Affirmed that a lot of QR code reader loads URL automatically upon scanning, hence putting the users at risk of phishing attacks and malware exploit. Another drawback of this solution is that most users ignore security warning [6], also this solution cannot handle zero day attacks of QR-code as it heavily depend on Google safe browsing API and phishtank API , both of which can be manipulated by the attackers [10].

Another prominent work from the literature is that of [14] which adopted machine learning techniques in the detection of QR code phishing attacks using Naïve Bayes classification algorithm. Though its detection rate of 93.35% performance was encouraging, but can be improve by a better methodology [3]. Also a better classification can be achieved using automatic data features selection and learning than manual with few instances. This will improve the model ability to handle zero day QR code phishing attacks.

REFERENCES

[1] Yao, H., & Shin, D. (2013). Towards preventing QR code based attacks on android phone using security warnings. *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security - ASIA CCS '13*, 341. https://doi.org/10.1145/2484313.2484357

[2] Kharraz, A., Kirda, E., Robertson, W., Balzarotti, D., & Francillon, A. A. (2014). Optical delusions: A study of malicious QR codes in the wild. *Proceedings - 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2014,* (December), 192–203. https://doi.org/10.1109/DSN.2014.103

[3] Krombholz, K., Frühwirt, P., Kieseberg, P., Kapsalis, I., Huber, M., & Weippl, E. (2014). QR code security: A survey of attacks and challenges for usable security. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8533 LNCS,* 79–90. https://doi.org/10.1007/978-3-319-07620-1_8

[4] Bode,G J. (2017). Quick-Response Codes and their acceptance in mobile shopping, 1–13. Proceedings of 9th IBA Bachelor Thesis Conference, July 5th, **2017**, Enschede, The Netherlands. https://essay.utwente.nl/72717/

[5] Dudheria, R. (2017). Evaluating Features and Effectiveness of Secure QR Code Scanners. In *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)* (pp. 40–49). https://doi.org/10.1109/CyberC.2017.23

[6] Shaikh, A. N., Shabut, A. M., & Hossain, M. A. (2017). A literature review on phishing crime, prevention review and investigation of gaps. *SKIMA 2016 - 2016 10th International Conference on Software, Knowledge, Information Management and Applications,* 9–15. https://doi.org/10.1109/SKIMA.2016.7916190

[7] Tao, L. (2017). QR code scams rise in China, putting e-payment security in spotlight | South China Morning Post. *South China Morning Post*. Retrieved from http://www.scmp.com/business/china business/article/2080841/rise-qr-code-scams-china-puts-Online-payment-security

[8] Sharma, V. (2012). A study of malicious qr codes, *3*(5), 3–8. Retrievedfromhttps://www.academia.edu/1864785/A_Study _of_Malicious_QR_Codes

[9] Thompson, Nik and Lee, Kevin (2013) "Information Security Challenge of QR Codes," *Journal of Digital Forensics, Security and Law*: Vol. 8 : No. 2 , Article 2.

DOI: https://doi.org/10.15394/jdfsl.2013.1143

[10] Lerner, A., Saxena, A., Ouimet, K., Turley, B., Vance, A., Kohno, T., & Roesner, F. (2015). Analyzing the Use of Quick Response Codes in the Wild. *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '15*, 359–374. https://doi.org/10.1145/2742647.2742650

[11] Targeted, F. (2018). Cyren Cyber threat, (April).Retrieved from:https://evessio.s3.amazonaws.com/customer/8c4659ee -526a-4e9c-89dc-f6f4c3c1a789/event/ipexpo europe/2018-Exhibitors/cyren-1_Cyren_Phishing.pdf

[12] T. Ishihara and M. Niimi, "Compatible 2D-Code Having Tamper Detection System with QR-Code," *2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kitakyushu, 2014, pp. 493-496. doi: 10.1109/IIH-MSP.2014.129

[13] Banu, M. N., & Banu, S. M. (2013). A Comprehensive Study of Phishing Attacks. *International Journal of Computer Science and Information Technologies, 4*(6), 783–786. Retrievedfromhttp://citeseerx.ist.psu.edu/viewdoc/download ?doi=10.1.1.643.766&rep=rep1&type=pdf

[14] Alnajjar, A., Sains, U., Manickam, S., Sains, U., Elejla, O., & Sains, U. (2016). QRphish: An Automated QR Code Phishing Detection Approach, (August). https://doi.org/10.3923/jeasci.2016.553.560