

QR 코드 복구율 개선을 위한 연구

김홍준¹, 김윤상², 이명재³, 노인기⁴, 유동영⁵^{1,2,3,4} 홍익대학교 소프트웨어융합학과 학부생⁵ 홍익대학교 소프트웨어융합학과 교수Was56475@gmail.com, kim31899@naver.com, audwolego@naver.com, revol083@naver.com, ydy@hongik.ac.kr

A Study To Improve Recovery Rate Of QR Code

Hong-Jun Kim¹, Yun-Sang Kim², Myeong-Jae Lee³, In-Ge No⁴, Dong-Young Yoo⁵^{1,2,3,4,5}Dept. of Software and Communication Engineering, Hongik University

요 약

QR 코드는 정부에서 지원하는 자전거 사업, wifi, 장비 및 공공기관 등의 관리 등 여러 분야에서 실생활과 밀접하게 쓰이고 있다. 하지만 QR 코드는 대체적으로 외부에 노출되어 있으며 이로 인한 훼손, 변조 등의 위험이 있다. 해당 연구에서는 위에 제시된 문제점을 개선하기 위해 플랫폼마다의 QR 코드의 규칙을 찾아내어 변조에 대한 위험을 줄이고 복구율을 높여 훼손에 대한 허용 범위 등의 개선을 기대한다.

1. 서론

QR 코드는 2020년 기준 QR 코드 시장의 확대를 위한 정부의 적극적인 지지가 예정되어 있었으며 이에 따라 QR 코드 잠재성장률이 매우 좋아 보였다[1]. 그리고 2024년 기준 정부 지원 자전거 사업, 장비 및 공공기관 관리, wifi, 신분 확인 등 여러 분야에서 실생활과 밀접하게 존재한다 이러한 QR 코드는 대체적으로 외부에 노출되며 이를 악용한 훼손이나 변조 등이 있을 수 있다. 이 때문에 훼손 시에 QR 코드 교체 비용 및 변조로 인한 피해 발생 등이 문제가 된다. 위에 제시된 문제들을 해결하기 위해 플랫폼마다 각각의 QR 코드의 규칙이 있는 것을 이용하여 QR 코드가 규칙에서 벗어났을 때는 변조에 대한 대처를 가능하게 한다. 또한 QR 코드가 많이 훼손되어도 규칙에 따라 어느정도 복구를 할 수 있기 때문에 교체 비용도 더욱 절감할 수 있을 것으로 예상된다. 이런 기대할 점들이 해당 연구를 진행하기 위한 근거가 될 수 있다고 생각한다.

2. QR 코드의 배경

QR 코드는 2차원 바코드 중의 일부이며 1차원 바코드에서 더욱 많은 정보를 저장하기 위해 고안됐으며[2] 이에 따른 1차원 바코드와의 차이점은 아래 표와 같다.

<표 1> 2차원 바코드와 1차원 바코드 비교

구 분	2차원 바코드	1차원 바코드
데이터 형식	영문, 숫자, 기호, 사진, 음성, 지문, 서명 등	영문, 숫자, 기호
데이터 용량	약 2,000바이트	약 20바이트
데이터 밀도	고밀도	저밀도
심볼 크기	정사각형(최소화 가능)	정보량이 커질수록 심볼이 커짐
판독 속도	정보량에 따라 속도에 영향	판독속도 양호
판독 방향	한방향 또는 전방향	한방향
DB 연계	불필요(직접코드)	필요
데이터 암호화	가능	불가
전자서명 인증	가능	불가
오류 검출	가능	가능
훼손 경정	가능	불가

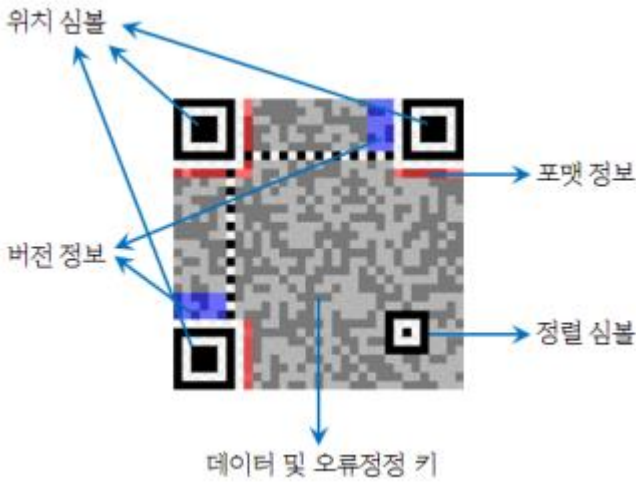
또한 QR 코드는 이러한 점 덕분에 국제 표준이 되기도 하였다. 참고로 여기서 선정된 국제 표준 2차원 바코드는 4가지가 있으며 각각 미국 심보로지의 PDF-147, 미국 ID Matrix의 Data Matrix, 미국 UPS의 MaxiCode, 일본 텐소 웨이브의 QR Code가 있다.[2] 아래 표는 국제표준 2차원 바코드 특징 및 차이점을 나타낸다.

<표 2> 국제표준 2 차원 바코드 특징 및 차이점

구분	QR Code	PDF-147	Data Matrix	MaxiCode
심볼				
개발업체	일본 덴소웨이브	미국 심브로지	미국 ID Matrix	미국 UPS
코드유형	매트릭스 방식	스택 방식	매트릭스 방식	매트릭스 방식
정보의 유형	영문 및 숫자, 한글 도형, 화상, 아스키(128문자)2진 데이터	영문 및 숫자, 한글 도형, 화상, 아스키(128문자)2진 데이터	영문 및 숫자, 한글 도형, 화상, 아스키(128문자)2진 데이터	아스키(128 문자), 2진 데이터
최대 정보량	숫자 7,089 영문숫자 4,464 2진 3,096바이트	숫자 2,725 영문숫자 1,850 2진 1,108 바이트	숫자 3,116 영문숫자 2,335 2진 1,566 바이트	숫자 138 영문숫자 93
오류정정 기능	7%, 15%, 25%,30%의 4단계	0-8단계, 최대 80% 정도에서도 판독 가능	28%-62.5% (EOC200의 경우)	25%, 50%의 2단계
특징	심볼 모서리 각에 3개의 파인더에 의한 고속판독 가능. 레이저 또는 이미지 판독 방법	1차원 바코드 판독기로 판독 가능, 2차원 시장 심볼 중에서 시장의 70% 이상 점유	정보화 밀도가 매우 높음. 심볼의 극소화가 가능하고 레이저 또는 이미지 판독 방법	심볼 한 가운데의 3종의 파인더에 의한 고속판독 가능. 심볼의 크기는 고정. 레이저 또는 이미지 판독 방법
이용 분야	부품관리, 자동차 등의 제조물류 분야	행정 및 군사용, 유통서비스, 제조, 인터넷 투표	액정 PCB, IC칩, 메모리 등의 제조 분야와 인터넷 투표, 추적등의 물류분야	물품관리, 구분, 추적 등의 물류분야
ISO/IEC의 의 표준화	AIMI, JIS, JEIDA 등	AIMI, ANSI, AIAG, EIA, AFMA, USD OD, UPU 등	AIMI, ANSI, SEIM, EIA, AIAG, UPU 등	AIMI, ANSI, EIA, AIAG 등

3. QR 코드의 구성

QR 코드는 위치 심볼, 정렬 심볼, 버전 정보, 포맷 정보와 함께 데이터 및 오류정정 키로 구성[3]되어 있으며 아래 그림과 같다.



(그림 1) QR 코드의 구성

또한 QR 코드는 저장 용량에 따라 21x21 크기의 버전 1 부터 177x177 크기의 버전 40 까지 40 개의 버전으로 구성 되어 있다.

4. QR 코드의 해석

QR 코드의 구성을 알면 해석을 할 수 있으며 아래에 제시 되어 있는 알고리즘을 순서대로 진행하면 된다. 각각 처음부터 말하자면 버전의 수, 레벨, 오류 정정 레벨 등 포맷 정보를 확인, 마스크 패턴으로부터 원래 코드 복구, 복구된 코드로부터 데이터 문자열 찾기, 오류 정정을 수행

하고 코드 꼬리 부분 제거, 코드, 워드로부터 근본 데이터를 찾기 순[4]이다. 아래는 해당 알고리즘을 표시한 그림이다.

1. 버전의 수, 레벨, 오류정정레벨 등 포맷 정보를 확인 한다.
2. 마스크 패턴으로부터 원래의 코드를 복구한다.
3. 복구된 코드로부터 데이터 문자열을 찾아낸다.
4. 오류정정을 수행하고 코드 꼬리 부분을 제거한다.
5. 코드 워드로부터 근본 데이터를 찾아낸다.

(그림 2) QR 코드 해석 알고리즘

5. QR 코드의 Error Correction

기본적으로 QR 코드는 데이터를 매핑하고 남은 공간에는 훼손에도 복구할 수 있는 Error Correction 을 남긴다. 해당 Error Correction 은 주로 Reed-Solomon code 를 사용 [5]하며 해당 기법으로 일부 데이터들을 복구를 진행한다.

6. 연구의 방향성

해당 연구에서는 위에서 제시된 지식을 이용하여 복구율을 높이는 것이 목표이다. 이에 따라 플랫폼마다의 QR 코드의 규칙을 찾는 것이 첫 번째 과제일 것이며 해당 규칙은 QR 코드를 해독하면서 찾아야 할 것이다. 그리고 해당 규칙을 찾았다면 CNN 이나 NIN 등 딥러닝 및 머신러닝을 적용시켜 규칙의 정확성을 증가시키고 QR 코드의 복구 방법을 다양화할 것이다. 또한 Spam QR 코드를 머신러닝이 규칙의 벗어난 정도를 확인하여 Spam 인지 아닌지 확인하는 기능도 검증해 볼 것이다.

참고문헌

- [1] 이은지 장지영, “QR Code 관련 연구 동향 분석”, 『한국컴퓨터정보학회 하계학술대회 논문집』, 제 29 권, 제 2 호, 2021 년, p.367-368
- [2] 박은지, “사례 분석을 통한 QR 코드의 활용방안 연구 -물류 유통분야를 중심으로-”, 『한국환경경제학회지』, 제 29 권, 제 4 호, 2013 년, p.247-264
- [3] 양형규, “QR 코드의 보안 취약점과 대응 방안 연구” 『한국인터넷방송통신학회 논문지』, 제 12 권, 제 1 호, 2012 년, p.83-89
- [4] 박찬정 현정석, ”스마트폰 기반의 QR 코드 해석기 성능분석 및 응용개발”, 『한국해양정보통신학회 논문지』, 제 13 권, 제 10 호, 2009 년, p.2242-2250
- [5] Matt Ket, “Reed-Solomon Error-correcting Codes The Deep Hole Problem”, 『UC Irvine Electronic Theses and Dissertations』, ProQuest ID: Ket_uci_0030D_13409. Merritt ID: ark:/13030/m541948m., 2015 년