정 보 보 호 학 회 지
제27권 제1호, 2017. 2

보안성 및 사용성 측면에서의 CAPTCHA 동향

조 금 환*, 최 주 섭*, 김 형 식**

요 약

웹 사이트에서 자동화 공격 도구를 이용한 다양한 종류의 공격을 방지하기 위한 보안 솔루션으로 CAPTCHA가 널리 이용되고 있다. 그러나 동시에 CAPTCHA를 해결하는 자동화 도구에 대한 연구가 진행되면서 CAPTCHA에 사용되는 텍스트 이미지(예: 숫자, 글자)를 더욱 어렵게 만들게 되었다. 그 결과 사용자도 CAPTCHA를 해결하는데 어려움을 겪게 되었고, 결론적으로 보안성을 높이기 위해 사용성을 감소시킨 결과를 초래 하였다. 본 논문에서는 텍스트, 오디오 및 이미지 기반 CAPTCHA로 분류하여 보안성과 사용성 측면에서 분석하고자 한다.

Ⅰ. 서 론

최근 컴퓨팅 성능이 우수한 머신들을 이용한 자동화 공격이 빈번하게 발생하고 있다. 이러한 자동화 공격을 이용해서 웹사이트에 대한 DDoS 공격이나 스팸 메시지를 대량으로 유포하는 등의 공격들이 가능해지면서 이를 방지하기 위한 보안 솔루션들이 연구되고 있다. CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)는 시도 응답(challenge-response) 테스트로 웹사이트를 사용하는 사용자가 인간인지 여부를 결정하기 위해 널리 사용되는 대표적인 자동화 공격 방지 보안 솔루션이다. 가장 많이 사용되는 CAPTCHA는 텍스트 기반으로 숫자나 글자로 구성된 이미지(*challenge*)가 주어지면 동일한 숫자나 글자를 입력(*response*)하는 방식이다. 그러나 출력된 텍스트 이미지를 자동으로 인식하는 기법(예: Optical Character Recognition)을 이용하여 CAPTCHA를 공격하는 다양한 공격들이 연구되었다 [3, 4, 5]. 이러한 CAPTCHA 공격에 대한 비용을 높이기 위해 텍스트 이미지(예: 숫자, 글자)를 회전시키거나(Rotation) 가로획을 이어 버리는(Arc) 등의 방법을 사용하였다. 결과적으로 봇(Bot)을 이용한 자동화 공격으로부터 저항력이 증가하여 보안성이 강화되었지만 웹사이트를 이용하는 사용자도 육안으로 구별하기 힘든 수준의 왜곡된 텍스트 이미지들로 인해 사용성이 크게 감소되는 단점이 동시에 발생하였다. 또한 오디오 및 이미지 기반 CAPTCHA를 이용하여 텍스트 기반 CAPTCHA의 문제점을 해결하려는 연구도 진행되었다 [7]. 그러나 오디오 및 이미지 기반 CAPTCHA의 경우에도 봇의 의한 자동화 공격을 방지하기 위해 사용하는 기법들로 인해 보안성과 사용성 측면에서 장점과 단점이 존재하고 있다.

본 논문에서는 최근 웹사이트에서 실제로 사용되는 CAPTCHA를 텍스트, 오디오 및 이미지 기반으로 분류하고 보안성과 사용성 측면에서 분석한다. 분류한 CAPTCHA는 다음과 같이 설명할 수 있다 [1].

- **텍스트 기반 CAPTCHA:** 정교하게 왜곡된 텍스트 이미지를 동일하게 입력하는 기법
- **오디오 기반 CAPTCHA:** 숫자 및 문자와 노이즈가 포함된 소리를 듣고 동일하게 입력하는 기법
- **이미지 기반 CAPTCHA:** 주어진 문제에 알맞은 이미지를 선택하는 기법

본 논문의 구성은 다음과 같다. 2장에서는 텍스트 기반 CAPTCHA에 대해 보안성 및 사용성 측면에서 분석한다. 오디오 기반 및 이미지 기반 CAPTCHA에 대해서 각각 3장과 4장에서 분석하고, 마지막으로 5장에서

본 연구는 한국연구재단 논문연구과제(95-0100-23-04-3) 지원 및 한국대학교 논문연구소 관리로 수행되었습니다.
* 성균관대학교 전자전기컴퓨터공학과 (geumhwan@skku.edu, cjs1992@skku.edu)
** 교신저자, 성균관대학교 전자전기컴퓨터공학과 (hyoung@skku.edu)

본 논문에 대한 결론을 도출할 것이다.

Ⅱ. 텍스트 기반 CAPTCHA

2.1. 텍스트 기반 CAPTCHA의 특징

텍스트 기반의 CAPTCHA는 보편적으로 가장 많이 사용되는 CAPTCHA 중의 하나로 왜곡된 텍스트 이미지를 복호화하고 입력함으로써 테스트를 통과할 수 있다[2]. 아래 [그림 1]과 같이 텍스트 기반 CAPTCHA는 *challenge*에 해당하는 왜곡된 텍스트 이미지가 주어지고 *response*로써 동일한 텍스트를 입력하는 방식이다.

또한 [그림 1]에서 볼 수 있듯이 왜곡된 텍스트 이미지(예: 글자를 휘고 가로획을 이어버림)를 사용함으로써 봇에 의한 자동화 공격에 대한 저항력을 증가시키는 방법을 사용한다. 이러한 텍스트 기반 CAPTCHA의 텍스트 이미지를 구성하는데 사용되는 기술은 다음과 같다.

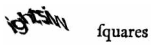
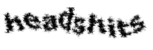
- **Connecting Characters Together (CCT)**기법은 텍스트 기반 CAPTCHA에서 가장 많이 사용되는 기법으로 1) 글자간 겹침(Overlap)은 허용되지만 노이즈 아크(Arc)는 없는 기법(가로획을 이어버리는 기법), 2) 노이즈 아크는 허용되지만 겹침이 없는 기법, 3) 노이즈 및 아크가 없는 기법의 3가지 종류로 생성된다. CCT 기법은 봇에 의한 글자 분할 및 인식 공격에 저항력을 갖고 있다. 그러나 최근 연구 결과 [3, 4, 5]에 의하면 CCT 기법은 자동화 공격에 취약한 것을 알 수 있다.
- **Hollow** 기법은 텍스트의 윤곽선만 표시하는 형태로 모든 텍스트는 서로 연결되도록 설계되었다. 또한 보안성과 사용성을 동시에 향상시키는데 목적을 두고 개발되었다. CAPTCHA 자동화 공격에 기본적으로 이용되는 분할(Segmentation) 및 인식 (Recognition) 방법에 저항력이 높다.
- **Character Isolated** 기법은 앞서 언급한 기법들과는 달리 각각의 문자가 서로 독립적으로 나타난다. 문자들이 서로 독립적으로 표현되는 대신 각각의 문자에 대한 왜곡이 다른 기법에 비해 심한 편이다.

〔그림 1〕 텍스트 기반 CAPTCHA(reCAPTCHA 1.0)의 예

대부분의 CAPTCHA는 위의 기법들을 사용하며 실제 웹사이트에서 사용되고 있는 CAPTCHA의 종류 및 샘플은 아래 [표 1]과 같고 각 CAPTCHA의 특징은 다음과 같다.

- **reCAPTCHA**는 Google(검색엔진), Facebook 및 Twitter(소셜 네트워크 서비스)에서 이용되고 있다. 텍스트 이미지로 숫자만 사용한다는 특징이 있고 길이, 크기를 변경할 수 있으며 회전이 가능하다. 또한 CAPTCHA를 구성하는데 CCT 기법을 사용한다.
- **Yahoo!**는 Yahoo의 웹사이트에 사용되며 hollow 기법을 사용한다. 텍스트 이미지에 대해 폰트, 회전, 왜곡 및 길이 변경이 가능하다.
- **Wikipedia**는 wikipedia.com의 웹사이트에서 사용되고 character isolated 기법을 사용한다. 텍스트 이미지의 길이를 변경할 수 있고, 숫자는 사용하지 않으며 문자만 사용한다는 특징이 있다.

〔표 1〕 웹사이트에서 사용되고 있는 텍스트 기반 CAPTCHA의 종류〔4〕

종류	reCAPTCHA	Yahoo!	Wikipedia	Microsoft	Amazon	ebay
웹사이트	google, facebook, youtube, linkedin, twitter, wordpress	yahoo.com, yahoo.co.jp	wikipedia.org	bing.com, MSN, Hotmail, Window Live	amazon.com	ebay.com
샘플						

〔표 2〕 DeCAPTCHA 종류 및 특징

종류	특징
GSA Captcha Breaker	윈도우 기반 소프트웨어로 600개의 CAPTCHA를 풀 수 있고 알고리즘 수정하거나 새로운 알고리즘을 추가할 수 있다.
DeCaptcha	OCR 기술을 사용하는 온라인 서비스로 8개 이상의 프로그램 언어에 대한 API를 지원한다.
Captcha Sniper	윈도우 애플리케이션으로 1,000개 이상의 CAPTCHA를 풀 수 있고, 매주 새로운 CAPTCHA에 대한 업데이트를 진행하고 있다.
DeathByCaptcha	인간이 해결하는 CAPTCHA와 OCR 시스템이 해결하는 CAPTCHA를 결합한 하이브리드 시스템으로 다양한 프로그램 언어에 대한 API를 지원한다.
BypassCaptcha	써드파티(third-party) 소프트웨어와 결합하기 쉬운 특징을 갖고 있고 자동화된 CAPTCHA 인지를 위한 다양한 디코딩 기능을 제공한다. 또한 다양한 프로그램 언어에 대한 API를 지원한다.
Image Typerz	가장 큰 특징은 풀기 어려운 CAPTCHA를 해결하지 않고, VIP 우선권을 갖고 있다면 10초 이내에 CAPTCHA를 해결할 수 있으며 추가 비용이 요구된다.
ExpertDecoders	인간 기반의 자동 CAPTCHA 해결 소프트웨어로 양질의 서비스를 제공한다. End-user의 경우 BypassCaptcha.com API나 De-Captcha API를 사용할 수 있으며, 소프트웨어 공급 업체나 개발자의 경우 PHP, C#등의 예제를 제공한다.
9kw.eu	대부분의 표준 CAPTCHA에 대해 30초 이내에 문제를 해결할 수 있고 인터페이스 및 플러그인의 사용이 용이하다.

- **Microsoft**는 주로 Microsoft에서 제작한 서비스(예: bing.com)에 사용되고 character isolated 기법을 기반으로 텍스트 이미지의 길이, 폰트 크기 및 회전이 가능하다.
- **Amazon**은 아마존닷컴에서 사용되고 CCT 기법을 사용한다. 텍스트 이미지에 대해 회전이 가능하며 특이한 점은 일정한 크기의 폰트를 사용한다.
- **ebay**는 ebay의 웹사이트에서 사용되고 CCT 기법을 사용한다. 폰트 크기를 변경할 수 있으며 회전 사용이 가능하다.

2.2. 텍스트 기반 CAPTCHA의 보안성

CAPTCHA 설계의 기본적인 원칙은 인간은 쉽게 *response* 할 수 있고 봇은 어렵게 하는데 있다. 2.1장에서 설명한 바와 같이 텍스트 기반 CAPTCHA는 다양한 기법을 이용하여 봇에 의한 자동화 공격을 방지하고 있다. 그럼에도 불구하고 텍스트 기반 CAPTCHA가 봇을 이용한 자동화 공격에 취약하다는 다양한 연구들이 진행되었다.

Yan 등 [3]은 봇을 이용하여 Microsoft CAPTCHA에 대한 자동화 공격을 진행하였다. 텍스트 이미지를 분할(Segmentation)하고 각각의 문자를 인식(Recognition)하는 방법으로 글자 분할의 성공률은 90% 이상, 평균 80ms의 시간 소요되었다는 연구결과를 발표하였다. 또한 전체 텍스트 이미지를 복원하는데 성공한 공격 성공률은 약 60%정도였다. 기본적인 CAPTCHA 설계의 목표는 자동화 공격을 시도했을 때 0.01% 이상 공격이 성공하면 실패한 시스템으로 간주하기 때문에 상당히 높은 수준의 공격 성공률을 나타냈다.

Bursztein 등 [5]은 실제 웹사이트에서 사용되고 있는 다양한 텍스트 기반 CAPTCHA를 공격하기 위해 기계학습(Machine Learning)을 이용하였다. 기계학습을 통해 분할과 인식을 동시에 수행함으로써 텍스트 이미지의 정보와 문맥을 추출한다. 제안하는 기법을 이용하여 Baidu에 대해 38.68%(2011년 버전) 및 55.22%(2013년 버전), CNN에 대해 51.09%, eBay에 대해 51.39%, Wikipedia에 대해 28.29%, 마지막으로 Yahoo에 대해 5.33%의 성공률을 보였다. 그러나 reCAPTCHA에 대해 22.67%(2011년 버전) 및 22.34%(2013년 버전)로 어려운 CAPTCHA에 대한 공격 성공률이 상대적으로 낮았다.

Gao 등 [4]은 이러한 문제점을 해결하기 위해 새로운 공격 기법을 제안하였다. Log-Gabor 필터를 이용하여 문자의 요소들(Components)을 추출하고 인식 엔진(k-Nearest Neighbors)을 이용하여 인접한 문자의 요소

〔표 3〕 사용성 유저스터디에 사용된 CAPTCHA의 종류 〔6〕

종류	Normal Text	Blot Mask	Line Mask	Thread Noise	Global Warp	Geometry Noise
샘플						

의 다양한 조합을 시도한 후 가장 올바른 조합으로 결과를 선택한다. 제안한 기법을 이용해서 reCAPTCHA에 대해 공격을 진행한 결과 77.2%의 성공률과 평균 10.27초가 공격에 소요되었다.

CAPTCHA를 공격하기 위한 연구 뿐 아니라 상용화 제품이나 무료 소프트웨어들이 개발되었다([표 2] 참조). 플랫폼에 상관없이 CAPTCHA를 해결할 수 있는 소프트웨어도 다수 존재하기 때문에 실제 웹사이트에서 사용되는 텍스트 기반 CAPTCHA는 자동화 공격에 취약하다고 볼 수 있다. 또한 더 높은 성공률과 빠른 시간 내에 텍스트 기반 CAPTCHA를 공격하기 위한 연구가 더욱 지속될 것으로 예상된다.

2.3. 텍스트 기반 CAPTCHA의 사용성 분석

앞서 언급한 바와 같이 자동화 공격에 취약한 CAPTCHA의 단점을 극복하기 위해 다양한 기법들 적용되었다. 예를 들어 왜곡된 텍스트 이미지를 더욱 왜곡시키는 방법을 사용하거나 문자의 회전을 많이 주는 등의 방법이 사용되었다. 그러나 이러한 방법들은 자동화 공격에 대한 저항력을 증가시키는 반면에 사용자도 CAPTCHA를 해결하는데 어려움이 발생하는 사용성에 문제가 생겼다. 결론적으로 사용성이 크게 감소되면서 사용자가 웹사이트를 사용하는데 불편함을 초래하였다.

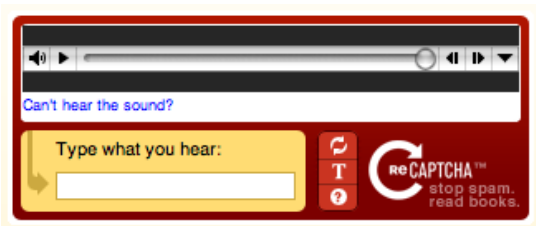
Lee 등 [6]은 연령으로 구분한 2개의 그룹을 만들고 텍스트의 찌그러짐 정도에 따른 태스크(task)를 24명의 참가자들에게 부여하는 유저스터디를 실시하였다. 23-24세 12명과 50-56세 12명을 대상으로 실험을 진행하였다. 6개의 세션 중 [표 3]의 Normal Text가 항상 우선적으로 할당되었고 나머지 5개의 CAPTCHA는 랜덤하게 할당되었다. 50-56세 그룹의 참가자들이 23-24세 그룹의 참가자들에 비해 응답시간이 더 오래 걸렸고 실패 확률도 더 높은 것으로 나타났다. 결론적으로 동일한 CAPTCHA에 대해 연령별로 인지하는 능력이 다른 것을 확인할 수 있었다. 향후 CAPTCHA를 설계할 때 위와 같은 요소들을 고려해서 설계할 필요가 있다.

Ⅲ. 오디오 기반 CAPTCHA

3.1. 오디오 기반 CAPTCHA의 특징

오디오 기반의 CAPTCHA는 시각 장애인 등의 텍스트 기반의 CAPTCHA를 구별하기 힘들 경우 보조수단으로 활용하거나 인터넷 전화를 통해 스팸전화를 거는 SPam over Internet Telephony (SPIT)를 막는 시스템에 사용된다 [9]. 오디오 기반 CAPTCHA는 *challenge*에 노이즈가 포함된 음성이 주어지고 *response*로써 주어진 음성을 입력하거나 연산 결과(예: 덧셈, 뺄셈 등)를 입력하는 방식이다. 아래 [그림 2]는 오디오 기반의 CAPTCHA 중 하나인 reCAPTCHA의 오디오 버전이다. 재생 버튼을 누르면 주변 노이즈와 함께 알파벳 혹은 숫자를 읽는 소리가 나온다. 주변 노이즈는 알아들을 수 없을 정도이기 때문에, 정확히 들리는 알파벳이나 숫자를 정확하게 입력하면 테스트를 통과할 수 있다.

오디오 기반 CAPTCHA에 사용되는 음성생성 기술은 일반적으로 *challenge*에 해당하는 소리와 사람은 알아듣기 힘든 노이즈를 동시에 들려주는 방법을 사용한다. 이러한 노이즈의 경우 사람은 알아듣기 힘들고 봇은 인지하기 쉽게 때문에 자동화 공격을 더 어렵게 만들 수 있는 중요한 특징이 될 수 있으며 오디오 기반 CAPTCHA에서 높은 수준의 보안성을 보장할 수 있는 요소이다 [10]. 그 외에 문자와 문자 사이에 간격을 무작위로 설정함으로써 자동화 공격을 어렵게 하는 방법들이 사용 된다.



〔그림 2〕 오디오 기반 CAPTCHA(reCAPTCHA)

아래의 [표 4]는 실제 웹사이트에서 사용 중인 오디오 기반 CAPTCHA의 종류 및 특징들이다 [9]. 각 CAPTCHA의 특징은 다음과 같다.

- **Google**은 0부터 9의 숫자로만 오디오 CAPTCHA가 구성되며, 길이는 5~10자로 고정되어 있지 않다. 그리고 한 사람의 목소리로만 CAPTCHA가 구성되어 있지 않고 성별 및 나이 등등이 서로 다른 사람들의 목소리들로 구성되어 있다. 한번 CAPTCHA에 통과하는 시간은 평균 10~15초 정도가 소요된다.
- **MSN**은 0부터 9의 숫자로만 오디오 기반 CAPTCHA가 구성되며 Google과 달리 길이는 10자로 고정되어 있다. 목소리는 한 명의 목소리로 통일되어 있으며, 한번 CAPTCHA에 통과하는 시간은 평균 5~9초 정도가 소요된다.
- **reCAPTCHA**는 문장 위주로 CAPTCHA가 구성되어 있으며 여러 사람의 목소리가 섞여있다. 통과하는데 걸리는 시간은 4초 이내로 다른 CAPTCHA에 비해 상대적으로 적은 시간이 소요된다.
- **ebay**는 0부터 9의 숫자로만 오디오가 구성되어 있으며, 길이 또한 6자로 고정되어 있다. 한 사람의 목소리로 통일되어 있지 않고 여러 사람의 목소리가 섞여서 구성된다. 평균 통과 시간은 4초 이내이며, reCAPTCHA와 동일하게 통과 시간이 빠른 편이다.
- **AOL**은 A~Z, a~z의 영어 알파벳과 0부터 9의 숫자로만 구성되어 있으며 길이는 8자로 고정되어 있다. 한 사람의 목소리로 통일되어 있지 않고 여러 사람의 목소리가 섞여서 구성된다. 평균 통과 시간은 10초 정도이다.
- **Digg**는 AOL과 마찬가지로 A~Z, a~z의 영어 알파벳과 0부터 9의 숫자로만 구성되어 있으며, 길이는 5자로 고정되어 있다. 여러 사람의 목소리를 사용하지 않고 한 사람의 목소리로 구성되었으며, 평균 통과 시간은 8초 정도이다.

오디오 기반 CAPTCHA는 음성을 듣고 텍스트를 입력하기 때문에 평균적으로 소요시간이 텍스트 기반 CAPTCHA보다 오래 걸리는 경향이 있다. 또한 봇에 의한 자동화 공격의 비용을 증가시키기 위해 여러 사람의 목소리를 이용한다는 특징이 있다.

3.2. 오디오 기반 CAPTCHA의 보안성

Tam 등 [11]은 classification 기술을 이용하여 Google, Digg, reCAPTCHA에 대한 자동화 공격을 진행하였다. Google의 오디오 기반 CAPTCHA의 경우 MFCC를 포함한 5가지의 추출법을 통해 오디오의 특징들을 추출한 뒤, AdaBoost, SVM, k-NN classifier를 사용하여 실험한 결과, 최대 67%의 공격 성공률을 보였다. 또한, Digg와 reCAPTCHA의 경우 Google의 오디오 기반 CAPTCHA를 공격한 방법과 동일하게 실험한 결과, 각각 71%, 45%의 공격 성공률을 보였다. 텍스트 기반의 CAPTCHA와 마찬가지로 0.01% 이상 공격이 성공하면 실패한 시스템이기 때문에 이는 상당히 높은 수준의 공격 성공률로 볼 수 있다. 실제로 오디오 기반 CAPTCHA에서는 숫자 및 알파벳이 주로 사용되는데 선택할 수 있는 스페이스가 매우 좁기 때문에 기계학습으로 트레이닝 하는 스페이스가 줄어들어 공격 성공률이 높아지는 결과를 초래할 수 있다. 따라서 오디오에 들어가는 숫자 및 알파벳의 범위를 넓히는 방법이 요구될 것이다.

Bursztein 등 [12]은 오디오 파일에서 Discreet Fourier Transform (DFT)을 적용하여 에너지 spike를 추출하여 기계학습을 통해 eBay의 오디오 기반 CAPTCHA에 대한 자동화 공격을 진행하였다. DFT를 사용한 방법을 적용하여 공격을 진행하였을 때 기존 연

〔표 4〕 웹사이트에서 사용되고 있는 오디오 기반 CAPTCHA의 종류〔9〕

종류	Google	MSN	reCAPTCHA	ebay	AOL	Digg
언급되는 텍스트의 종류	숫자	숫자	문장	숫자	알파벳 및 숫자	알파벳 및 숫자
Audio 시간	10~15초	5~9초	4초 이내	4초 이내	10초	8초
예시	13494437	1958488374	To be or not to be	128576	a8dbc93e	bu85e

구 [11]의 성공률인 71%보다 더 높은 75%의 공격 성공률을 보였다. 그 외에도 음향처리 기술 및 기계학습의 발전으로 인하여 오디오 기반의 CAPTCHA는 앞으로 자동화 공격을 위한 연구가 지속될 것으로 예상되며 높은 성공률을 나타낼 것으로 예상된다. 따라서 기본적으로 선택할 수 있는 스페이스를 증가시킬 수 있는 연구가 요구된다.

3.3 오디오 기반 CAPTCHA의 사용성

앞서 언급한 바와 같이 자동화 공격에 취약한 CAPTCHA의 단점을 극복하기 위해 다양한 기법들이 적용되었다. 예를 들어 노이즈의 볼륨을 더욱 크게 하거나, 숫자나 알파벳을 불러주는 주기를 무작위로 배치하는 등의 방법 등이 사용되었다. 그러나 이러한 방법들은 자동화 공격에 대한 안전성을 향상시키는 반면에 사용자가 CAPTCHA에서 제공되는 challenge를 해결하는데 어려움이 생겼다. 따라서 사용성이 크게 감소되면서 사용자가 웹사이트를 사용하는데 불편함을 초래하였다 [13].

Soupionis 등 [9]는 12개 정도의 오디오 기반의 CAPTCHA에 대해 분석을 진행하였다. 사람이 통과하는 성공률에 대해서는 영어가 모국어이며 주당 컴퓨터 사용시간이 20 시간 이상이 되는 참가자를 모집하여 각각 CAPTCHA 들의 *challenge*를 총 100개씩 시도함으로써 성공률을 실험을 진행하였다. 사람들이 가장 많이 사용하는 Google, MSN, reCAPTCHA에 대한 성공률은 각각 60%, 80%, 50%로 매우 낮은 결과를 나타냈다. 실제로 이는 3.2 장에서 언급한 자동화 공격에 대한 성공률과 비슷한 수준으로 CAPTCHA의 기본 설계 목적인 사람과 기계를 구분하고자하는 목적과는 부합하지 않는다. 또한 오디오 기반의 CAPTCHA는 텍스트 기반의 CAPTCHA에 비해 걸리는 시간이 오래 걸린다는 점에서 사용성이 매우 떨어지는 것을 확인할 수 있다. 최종적으로 컴퓨터를 능숙히 사용하는 사용자조차도 성공률이 높지 않은 것을 확인할 수 있었고 향후 CAPTCHA 설계 시 사용성 분석 등의 요소들을 고려해서 설계할 필요가 있다.

Ⅳ. 이미지 기반 CAPTCHA

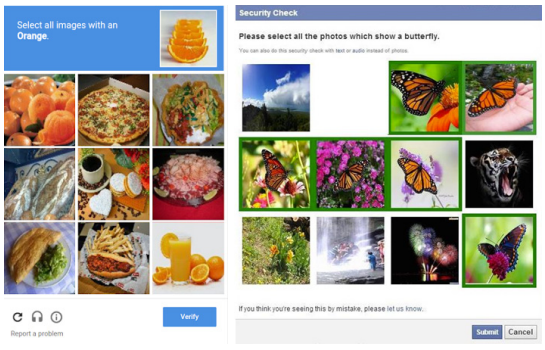
4.1. 이미지 기반 CAPTCHA의 특징

웹사이트에서 가장 많이 사용되는 CAPTCHA는 텍스트 기반 CAPTCHA이지만 반대로 이를 공격하기 위한 DeCAPTCHA도 많이 개발 및 연구되고 있다(2.1장 참조). 텍스트 및 오디오 기반 CAPTCHA의 문제점을 해결하기 위해 이미지 기반 CAPTCHA가 대안이 될 수 있다.

안전한 CAPTCHA 시스템을 위해 이미지 기반 CAPTCHA에 대한 연구가 진행되었다 [7]. 제안하는 시스템은 보안 수준에 따라 3×3(9개 조각)부터 5×5(25개 조각)로 분할된 퍼즐이 주어지고 그 중 오직 2개의 조각만 잘못된 위치에 주어진다. 잘못된 위치의 조각을 이동시켜 원래의 그림을 완성하면 CAPTCHA 테스트를 통과하는 방식이다.

그 외에 실제 웹사이트에서 사용되는 Google (reCAPTCHA 2.0)과 Facebook에서 사용하는 이미지 CAPTCHA가 가장 대표적이다. 아래 [그림 3]에서 볼 수 있듯이 주어진 이미지에서 문제로 주어진 이미지를 모두 선택하면 CAPTCHA를 통과할 수 있는 방식이다.

Google에서 사용하고 있는 reCAPTCHA 2.0의 경우 이미지 기반 CAPTCHA가 즉각적으로 주어지지 않는다는 특징이 있다. 아래 [그림 4]와 같이 “I’m not a robot”이라는 문구를 보고 클릭을 하면 웹페이지의 캐시나 마우스 움직임 등의 정보를 이용하여 봇인지 여부를 판단하고 추가적인 판단이 요구될 때 [그림 3]의 이미지 기반 CAPTCHA가 문제로 주어진다. 그러나 이미지 기반 CAPTCHA의 문제를 해결했음에도 불구하고



〔그림 3〕 Google(좌), Facebook(우) CAPTCHA의 예

여전히 의심스러운 경우에는 최종적으로 텍스트 기반 CAPTCHA가 주어진다. Google에서 사용하는 reCAPTCHA 2.0은 한 종류의 CAPTCHA를 이용했을 때 발생할 수 있는 단점들을 극복하기 위해 여러 종류의 CAPTCHA를 조합해서 사용하는 시스템이다.

4.2. 이미지 기반 CAPTCHA의 보안성

Gao 등 [6]은 제안하는 이미지 기반 CAPTCHA인 *Jigsaw puzzle*의 보안성을 평가하였다. N×N의 퍼즐 이미지의 경우 선택 가능한 모든 경우의 수가 N^2 이 된다 (2개의 조각만 이동시키면 됨). 무작위 공격 (Brute Force Attack)을 시도할 경우 모든 조합 가능한 경우의 수는 C_N^2가 된다. 따라서 단일 랜덤 추측의 경우의 수는 $1/C_N^2$가 된다. 또한 연구 결과에서 4×4 퍼즐의 경우 공격 성공률이 0.83%로 100번 시도했을 경우 1번미만으로 성공할 수 있기 때문에 상대적으로 텍스트 기반 CAPTCHA에 비해 보안성 측면에서 우수하다고 볼 수 있다.

Sivakorn 등 [8]은 딥러닝(Deep Learning) 기술을 이용하여 Google 및 Facebook의 이미지 기반 CAPTCHA에 대한 자동화 공격을 시도하였다. Google의 reCAPTCHA 2.0에서 봇을 판단하는데 사용되는 정보(캐시, 마우스 움직임 등)를 이용하였고 오픈소스로 공개된 다양한 딥러닝 알고리즘을 이용하여 reCAPTCHA 2.0에 대한 공격을 시도하였다. 연구 결과에 의하면 reCAPTCHA 2.0에 대해 70.78%의 공격 성공률과 19초 정도의 시간이 소요되었고, Facebook의 이미지 기반 CAPTCHA에 대해 83.5%의 성공률을 얻을 수 있었다. 전체적인 성공률을 비교해보면 텍스트 기반 CAPTCHA에 비해 자동화 공격에 더 취약하다는 것을 알 수 있다. 따라서 실제 웹사이트에서 사용되는 CAPTCHA의 보안성 향상 연구가 요구된다.

4.3. 이미지 기반 CAPTCHA의 사용성

Jigsaw puzzle [7]의 사용성을 평가하기 위해 250명의 사용자를 대상으로 유저스터디를 진행 하였다. 3×3 퍼즐의 경우 사용자에 의한 성공률이 86.7%를 보였다. 4×4 퍼즐의 경우는 85.5%, 5×5 퍼즐의 경우 86%의 성공률을 보였다. 모든 퍼즐에 대해서 약 15%정도의 사용자는 CAPTCHA를 통과하지 못했기 때문에 실제 시스템에 적용하기 위해서는 사용성에 대한 개선이 요구된다.

한편 Google이나 Facebook에서 사용되는 이미지 기반 CAPTCHA의 경우 현재까지 사용성에 대해 평가한 연구결과가 존재하지 않는다. 실제 웹사이트에서 사용되고 있는 CAPTCHA인 만큼 사용자들의 사용성 평가가 이루어져야 할 것이다.

Ⅴ. 결 론

본 논문에서는 웹 사이트에서 자동화 공격을 방어하기 위해 사용되는 보안 솔루션인 CAPTCHA에 대한 동향을 파악하였다. 텍스트, 오디오, 이미지 기반 CAPTCHA로 분류하여 보안성 및 사용성에 대한 분석을 진행하였다. 현재까지 제공되는 솔루션 중 보안성과 사용성을 동시에 만족하는 솔루션이 존재하지 않는다. 따라서 향후 CAPTCHA 설계 시 두 가지 고려사항을 만족할 수 있는 CAPTCHA의 개발이 요구된다.

참 고 문 헌

[1] Jeff Yan and Ahmad Salah El Ahmad, “Usability of CAPTCHAs Or usability issues in CAPTCHA design”, Symposium On Usable Privacy and Security, 2008.

[2] Jeff Yan and Ahmad Salah El Ahmad, “CAPTCHA Security: A Case Study”, IEEE Security & Privacy 7.4 (2009).

[3] Jeff Yan and Ahmad Salah El Ahmad, “A Low-cost Attack on Microsoft CAPTCHA”, ACM Conference on Computer and Communications Security, 2008.

[4] Haichang Gao, Jeff Yan, Fang Cao, Zhengya Zhang, Lei Lei, Mengyun Tang, Ping Zhang, Xin Zhou, Xuqin Wang and Jiawei Li, “A Simple Generic Attack on Text Captchas", Network and Distributed System Security Symposium, 2016.

[5] Elie Bursztein, Jonathan Aigrain, Angelika Moscicki, John C. Mitchell, “The End is Nigh: Generic Solving of Text-based CAPTCHAs”,

USENIX Workshop on Offensive Technologies, 2014.

[6] Ying-Lien Lee, Chih-Hsiang Hsu, “Usability study of text-based CAPTCHAs”, Displays 32 (2011).

[7] Haichang Gao, Dan Yao, Honggang Liu, Liming Wang, “A Novel Image Based CAPTCHA Using Jigsaw Puzzle”, IEEE Conference on Computational Science and Engineering, 2010.

[8] Suphannee Sivakorn, Iasonas Polakis, and Angelos D. Keromytis, “I AM Robot: (Deep) Learning to Break Semantic Image CAPTCHAs”, IEEE European Symposium on Security and Privacy, 2016.

[9] Soupionis, Yannis, and Dimitris Gritzalis, "Audio CAPTCHA: Existing solutions assessment and a new implementation for VoIP telephony", Computers & Security, pp.603-618, 2010.

[10] Jurafsky, Daniel, and James H. Martin, "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", Prentice-Hall, 2008.

[11] Tam, Jennifer, Simsa, Jiri, Hyde, Sean, and Von Ahn, Luis, "Breaking Audio CAPTCHAs.", Conference on Neural Information Processing Systems, 2008.

[12] Bursztein, Elie, and Steven Bethard, "Decaptcha: breaking 75% of eBay audio CAPTCHAs.", USENIX Workshop on Offensive Technologies, 2009.

[13] Bigham, Jeffrey P., and Anna C. Cavender., "Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use.", ACM Conference on Human Factors in Computing Systems, 2009.

〈저자소개〉

조 금 환 (Geumhwan Cho)
학생회원

2011년 2월 : 청주대학교 정보통신공학과 학사
2013년 2월 : 경희대학교 컴퓨공학과 석사
2014년 9월~현재 : 성균관대학교 전자전기컴퓨터공학과 박사과정
관심분야 : Usable security, 정보보호, 모바일 보안

최 주 섭 (Jusop Choi)
학생회원

2015년 8월 : 성균관대학교 컴퓨터공학과 학사
2015년 9월~현재 : 성균관대학교 전자전기컴퓨터공학과 석사과정
관심분야 : 정보보호, 디지털 포렌식, 역공학

김 형 식 (Hyoungshick Kim)
종신회원

1999년 2월 : 성균관대학교 정보공학부 학사
2001년 2월 : KAIST 컴퓨터 과학과 석사
2012년 2월 : University of Cambridge 컴퓨터공학과 박사
2013년 3월~현재 : 성균관대학교 전자전기컴퓨터공학과 조교수
관심분야 : 보안공학, 소셜 컴퓨팅