Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.27, NO.3, Jun. 2017 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2017.27.3.531

자동화 공격과 릴레이 공격에 저항하는 Emerging Image Cue CAPTCHA 연구*

양 원 석,† 권 태 경‡
연세대학교 정보보호연구실

Emerging Image Cue CAPTCHA Resisting Automated and Human-Solver-Based Attacks*

Wonseok Yang,† Taekyoung Kwon‡
Information Security Lab., Graduation School of Information, Yonsei University

요 약

CAPTCHA는 인터넷에서 서비스 혹은 자원을 요청하는 존재가 사람인지 아닌지를 구별하기 위한 테스트 기법이다. 이러한 대부분의 CAPTCHA들은 CAPTCHA 영상을 제3자에게 스트리밍 하여 제 3자가 사용자 대신 CAPTCHA에 응답하는 스트림 릴레이 공격에 의해 우회 될 수 있다는 문제점을 갖는다. 스트림 릴레이 공격에 저항성을 갖기 위해 인간의 인지 구조를 이용한 Emerging Image Game CAPTCHA가 제안되었으나, 사용성이 낮다는 문제점이 있다. 본 연구에서는 Emerging Image Game CAPTCHA의 사용성을 개선할 수 있는 Emerging Image Cue CAPTCHA 설계 방법과 CAPTCHA를 제안하며, 제안된 CAPTCHA에 대한 사용성 평가, 릴레이 공격 시뮬레이션을 통해 실질적인 개선이 이루어졌는지를 평가한다.

ABSTRACT

CAPTCHA is a verification scheme whether or not a human user has made a service request. Most CAPTCHAs that are based on text, image, or simple game suffer from vulnerability that can be compromised by automated attacks and stream relay attacks. To resist such attacks, CAPTCHA that utilizes human recognition as been suggested but it show poor usability for deploying in the Internet. We propose an Emerging Image Cue CAPTCHA that offers improved usability and resists stream relay attacks, as well. We also examine the usability of the proposed CAPTCHA and investigate the attack resistance by conducting user study and experiments on simulated network environment.

Keywords: CAPTCHA, Emerging Image, Automated attack, Static Relay Attack, Stream relay attack, Usability test

I. 서 론

1.1 연구 배경

Completely Automated Public Turing Test to Tell Computers and Humans Apart(이하 CAPTCHA)는 인터넷에서 서비스를 요청하는 존재가 사람인지 아닌지를 구별해낼 수 있는 보안 장치이다. 일반적인 텍스트 기반 CAPTCHA의 경우 OCR(Optical Character Recognition)기법 등을 이용하여 공격자가 쉽게 CAPTCHA 응답을 자동으로 입력하는 자동화 공격

Received(02. 27. 2017), Accepted(03. 27. 2017)
* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학 ICT연구센터육성 지원사업의 연구결과로 수행되었음 (IITP-2017-2012-0-00646). 또한 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.NRF-2015R1A2A2A01004792)
† 주저자, zmsenqn@yonsei.ac.kr
‡ 교신저자, taekyoung@yonsei.ac.kr(Corresponding author)

에 취약하다. 자동화 공격에 대응하기 위한 CAPTCHA인게임 기반 CAPTCHA의 경우, 공격자와 협력 관계에 있는 조력자가 원격지에서 화면 모니터링 및 CAPTCHA 응답을 공격자 대신 입력하는 릴레이 공격에 취약하다는 문제점을 갖고 있다. 또한 매우 단순하고 한정적인 패턴을 가진 게임 기반 CAPTCHA는 자동화 공격에조차 취약하다. 자동화 공격 및 릴레이 공격 모두에 대응하기 위한 방안으로 텍스트 기반 CAPTCHA의 텍스트에 대해 Emerging Image 기법을 사용한 CAPTCHA가 제시되었으나, 사용자들의 인식률이 저조하고 사용하기 불편하다는 한계점을 갖고 있으며 상용화에 실패하였다[1].

1.2 연구 목표와 범위

기존에 존재하는 텍스트 기반, 게임 기반 CAPTCHA는 대개 자동화 공격 및 릴레이 공격을 통해 우회할 수 있다는 문제점을 갖는다. 자동화 공격은 컴퓨터를 이용하여 CAPTCHA화면을 인식, 학습하여 CAPTCHA질의에 나타난 단어를 자동으로 입력한다거나 게임 조건을 자동으로 충족시켜 CAPTCHA를 우회하는 공격 기법을 의미한다. 릴레이 공격은 CAPTCHA 화면을 원격지에 있는 컴퓨터에 전송하고, 원격지의 컴퓨터를 보고 있는 제 3자가 CAPTCHA 질의에 응답하여 서비스 요청 주체가 직접 CAPTCHA에 응답하지 않고도 CAPTCHA를 우회 할 수 있는 공격 기법을 의미한다. 현존하는 텍스트 기반 CAPTCHA들은 컴퓨팅 성능의 발달, 공격 기법의 진화에 따라 자동화 공격에 취약하다는 문제점을 가지며, 자동화 공격에 대응하기 위한 게임 기반 CAPTCHA들은 화면을 원격지에 있는 제 3자에게 전송하여 제 3자가 실제 서비스 이용자 대신 CAPTCHA 질의에 응답하는 릴레이 공격에 취약하다는 문제점을 갖는다. 자동화 공격과 릴레이 공격에 모두 대응하기 위한 방안으로 Emerging Image 기법을 적용한 CAPTCHA가 이미 제안된 바 있으나[1], 사용성 평가에서 저조한 결과를 보이며 상용화에 실패한 사례가 있어 공격 저항성과 사용성을 모두 충족할 수 있는 연구가 필요하다.

본 연구의 목적은 세 가지로 첫째, 공격 저항성을 갖는 Emerging Image 기법을 적용한 CAPTCHA를 설계하는 것, 둘째, 설계한 CAPTCHA를 직접 구현하여 공격 저항성을 검증하는 것, 셋째, 사용성 및 인식률에 치명적인 영향을 주지 않는 CAPTCHA 구성요소에 Emerging Image 기법을 적용한 Emerging Image Cue CAPTCHA (이하 EIQ CAPTCHA)를 제안하고 이에 대한 사용성 평가를 진행하는 것이다.

II. 관련 연구

2.1 CAPTCHA

CAPTCHA는 인터넷에서 어떤 서비스를 요청한 주체가 사람인지 아닌지를 서버 측에서 판별할 때 쓰는 도구이다. 이 과정은 일방적인 확인 과정이 아닌, CAPTCHA 인터페이스와 서비스 요청 주체간의 상호작용을 통해 진행된다. CAPTCHA는 서비스 요청 주체에게 모종의 질문을 던지고, 질문을 받은 서비스 요청 주체는 해당 질문에 대한 응답을 제출한다. 제출한 응답이 서버 측에서 정답으로 판별되면 서비스 요청 주체를 사람으로 간주하여 서비스를 제공하고, 오답으로 판별된 경우에는 서비스를 거부한다. 1997년 인터넷 검색 엔진인 AltaVista에서 처음 서비스 되었으며, CAPTCHA는 다양한 기법과 형태로 발전하여 오늘날 대다수의 인터넷 기반 서비스에서 쉽게 접할 수 있게 되었다.

2.2 CAPTCHA의 기원

일반적으로 최초의 CAPTCHA는 1997년, 인터넷 검색 사이트인 AltaVista에서 서비스 되었다고 알려져 있다. 정확히 CAPTCHA라는 이름으로 사용되지는 않았으나, AltaVista 측 서버에 URL을 자동으로 제출하는 것을 방지하기 위한 보안 장치로 적용되었다. CAPTCHA라는 단어가 정의 된 것은 2003년으로, Luis von Ahn등이 학계에 발표한 것이 최초이다[2]. CAPTCHA의 필요성이 부각되었던 사례는 Luis von Ahn등의 논문에서 소개하고 있다. 1999년 11월에 있었던 최고의 컴퓨터 과학 대학원 과정을 뽑기 위한 온라인 투표 시스템에서 카네기멜론 대학(CMU)과 매사추세츠 공과대학(MIT)의 학생들이 자신들의 대학원 과정에 자동으로 투표하는 bot 프로그램을 개발하여 투표수를 조작한 사례가 바로 그것이다. 이 사례는 온라인 투표에서 투표할 수 있는 주체가 사람만이 아님을 보여주는 사례이다[1].

2.3 Emerging Image

Emerging Image는 2009년, N. J. Mitra 등이 제안한 이미지 처리 기법으로, 보기에 의미 없는 정보들을 취합하여 의미 있는 정보로 재인식하는 인간의 고유 인지 능력을 이용하는 이미지 처리 기법이다. Emerging Image는 배경정보가 있는 3D 오브젝트에 대해 주 오브젝트를 특정해내고, 주 오브젝트의 형태 및 윤곽선을 생략하여 그려내는 기법이다[15]. Emerging Image 렌더링 과정은 원본이 3D 오브젝트라는 점에서 응용성이 떨어진다는 한계를 갖는다.

이러한 Emerging Image의 3D 오브젝트 의존성을 해결하기 위한 방법으로 Yang, Cheng-Han 등은 사진으로부터 Emerging Image를 렌더링 하는 방법을 제안한다. 3D 오브젝트와는 달리 Splat의 기준이 없기 때문에 Superpixel Segmentation을 통해 Splat 과정을 보완하고, 사진으로부터 Edge를 탐지하여 Importance map을 구한 뒤, Superpixel Segmentation결과물과 Importance map의 정보를 합해 주 오브젝트의 Emerging Image 렌더링을 수행한다. 이후 배경을 채우기 위해 3D Emerging Image 렌더링에 사용된 copy-perturb-paste 기법을 사용하여 clutter를 생성한다. 주 오브젝트와 배경 clutter를 합치는 것으로 사진으로부터의 Emerging Image 렌더링 과정이 완료된다[16]. 사진으로 부터의 Emerging Image 렌더링은 3D 오브젝트로부터의 렌더링 보다 입체감은 떨어지나, 렌더링 가능한 대상이 무궁무진하게 많아진다는 데에 그 의의가 있으며, 기존 Emerging Image 렌더링 방식이 갖고 있던 한계점을 극복할 수 있다는 장점을 갖는다.

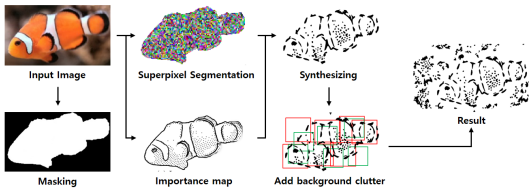


Fig. 1. Emerging Image Rendering from Photograph

2.4 Emerging Image 기법을 이용한 CAPTCHA

Song Gao 등은 Emerging Image 효과를 적용한 CAPTCHA를 제안하고, 제안한 CAPTCHA가 자동화 및 릴레이 공격에 저항할 수 있는지를 실험해 보았다. 이 논문에서 제안된 CAPTCHA는 Emerging Image Game CAPTCHA 이며, Emerging Image 효과를 CAPTCHA의 글자에 직접적으로 적용한 뒤, 노이즈의 배경에서 글자들이 움직이도록 하였다. 사용자는 이렇게 움직이는 글자들을 마우스로 드래그 하여 화면 왼쪽에 제시된 알맞은 장소에 끌어다 놓는 것으로 응답 할 수 있다. 이 CAPTCHA는 2D인 글자들에 대해 단순히 Emerging Image 효과를 주는 것에 그치지 않고, 처리된 글자들을 3차원 화 시키는 Pseudo 3D Visual Effect를 추가로 적용하였다. 이는 2차원 상의 오브젝트를 3차원 공간에 투영하는 기법으로 연속된 프레임들을 분석하여 글자의 윤곽을 수복하는 기법을 사용한 자동화 공격에 저항하기 위한 것이다. Song Gao 등은 제안한 CAPTCHA에 대해 자동화 및 릴레이 공격 저항성을 평가하였는데, 자동화 공격 기법으로는 density-based automated attack을 사용하였고, 릴레이 공격 실험에서는 인도-미국 간, 미국-미국 간 원격 공격 시뮬레이션 환경을 조성한 뒤 피험자들에게 이 환경에서 Emerging Image Game CAPTCHA를 풀어보도록 하였다. 실험 결과 자동화 공격 및 릴레이 공격 실험 모두에서 높은 공격 저항성을 보였으나, 사용성 평가 결과 50.9의 SUS(System Usability Scale)점수를 보여 사용에 불편하다는 결과를 얻었다[1].

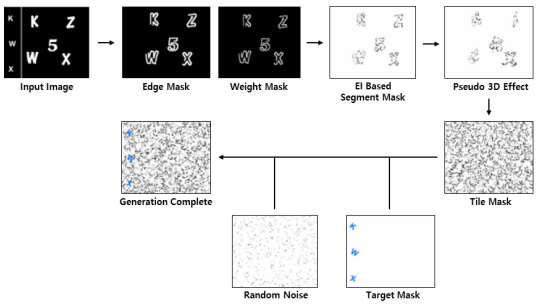


Fig. 2. Emerging Image Game CAPTCHA Construction

III. Emerging Image CAPTCHA 설계

3.1 위협 모델

CAPTCHA 시스템이 공격 받을 수 있는 모델은 크게 세 가지로 나눌 수 있다. 가장 먼저, CAPTCHA가 표시되고 있는 컴퓨터 자체에서 공격이 이루어지는 경우로, CAPTCHA를 풀어야 하는 사용자가 CAPTCHA에 대해 컴퓨터를 이용한 자동화 공격(Automated Attack)을 수행하는 경우이다. 공격 기법으로는 무작위 추측 공격, 기계학습을 이용한 공격, 이미지 처리 기법을 이용한 공격 등 다양한 공격 방식이 존재한다.

두 번째로, 사용자가 정지된 CAPTCHA 화면을 원격지의 제3자에게 전송하여 제3자가 CAPTCHA에 대한 해답을 다시 사용자에게 전송(Static Relay Attack)하여 CAPTCHA를 우회하는 공격 방법이 존재한다. 단순한 이미지 기반 CAPTCHA나 텍스트 기반 CAPTCHA는 이러한 공격에 의해 쉽게 우회될 수 있다는 취약점을 갖는다.

세 번째로, 사용자가 실시간으로 연속된 CAPTCHA 화면을 원격지의 제3자에게 전송하여 제3자가 CAPTCHA의 해답을 사용자에게 제공하거나, 사용자 대신 CAPTCHA 질의에 응답해주는 공격 방법이 존재한다(Stream Relay Attack). 이 공격 방법은 움직이는 텍스트 및 오브젝트를 이용한 CAPTCHA는 물론, 다소 복잡도가 있는 게임 기반 CAPTCHA를 우회할 수 있는 CAPTCHA에 대한 가장 강력한 공격이라 할 수 있다.

이러한 세 가지 위협 모델은 공격자가 CAPTCHA의 내부 구현 구조를 알 수 없고, CAPTCHA 질의에 대응하는 해답을 도청 및 해킹 등의 공격으로 미리 알거나 획득 할 수 없다는 전제에 기반을 둔다.

3.2 설계 개념 및 Emerging Image 기법의 적용 영역

EIQ CAPTCHA는 조작되지 않은 정상적인 이미지 혹은 영상에 Emerging Image 기법을 적용하여 사람은 그림의 내용을 파악할 수 있지만 기계 혹은 원격지의 사람에게는 내용을 파악하기 매우 어렵게 만드는 원리에 기반을 둔다.

본 논문에서 사용한 게임 기반 CAPTCHA 화면에 대해 Emerging Image효과를 적용하기 전과

Fig. 3. Roulette Game

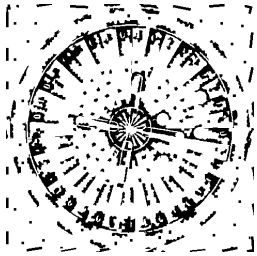


Fig. 4. Applying EI

후의 예시는 〔Figure 3, 4〕와 같다.

EIQ CAPTCHA는 기존에 제안되었던 Emerging Image Game CAPTCHA와 달리, 사용자가 인식 또는 조작해야 하는 부분에 대해선 Emerging Image 효과를 적용하지 않는다는 차별성을 갖는다. 실제로 Emerging Image Game CAPTCHA는 사용자가 읽어야 하는 글자들 자체에 Emerging Image 효과 및 Pseudo 3D effect를 적용하는데, EIQ CAPTCHA는 이와 달리 사용자가 조작해야 하는 부분 혹은 위치에 대한 실마리(Cue)에 대해서만 Emerging Image 효과를 적용한다는 차별성을 갖는다.

EIQ CAPTCHA는 HTML, Java Script, CSS를 사용한 480x480픽셀 해상도의 웹 룰렛 게임을 기반으로 한다. 이 룰렛 게임은 통상적인 룰렛과 달리 총 8개의 공을 사용한 변형된 룰렛 게임으로, 자동화 공격 저항성을 갖기 위한 것이다. 이 룰렛 게임에서 사용자가 인식해야 하는 단 하나의 공은 나머지 7개의 공들의 움직임과 구별되는 움직임을 갖는다. 사용자가 EIQ CAPTCHA에 응답하기 위해선 움직임이 다른 하나의 공을 식별해야 한다.

Emerging Image 효과가 적용된 룰렛 게임 영상을 제작하기 위해 원본 룰렛 게임 영상을 초당 30 프레임의 무압축 AVI파일로 녹화하고, 녹화한 영상의 모든 프레임에 대해 개별적으로 Emerging Image 효과를 적용하였다.

이후 Emerging Image효과가 적용된 모든 프레임들을 모아 다시 영상으로 재구성하여 EIQ CAPTCHA에 사용될 CAPTCHA 영상을 제작하였다.

Emerging Image 기법이 적용된 룰렛 영상은 C#언어로 개발한 GUI 프로그램에 탑재되었으며, 이 프로그램에서 룰렛 가장자리의 숫자를 룰렛 영상 위에 오버래핑 하여 사용자에게 표시하도록 하였다. 또한 이 프로그램은 실험 과정에서 CAPTCHA 응

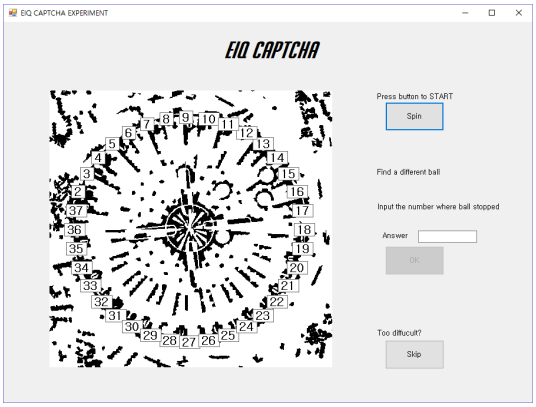


Fig. 5. EIQ CAPTCHA Program

답 소요 시간, 응답 시도 횟수, 최종 성공 여부, SUS 평가점수 등을 기록함으로써 사용성 평가에 필요한 평가 척도를 파일로 기록할 수 있도록 구현하였다. 시간은 밀리세컨(ms)단위로 측정되고 SUS 평가점수는 설문 과정에서 사용자의 응답에 따라 자동으로 계산되어 파일에 기록된다.

IV. 안정성 분석

4.1 자동화 공격 저항성

본 연구에서 실험에 사용한 룰렛 게임 기반 EIQ CAPTCHA는 자동화 공격에 저항하기 위하여 Emerging Image효과를 적용하는 것뿐만 아니라 움직이는 공을 여러 개 갖는다. 본 연구의 EIQ CAPTCHA는 사용자에게 실마리(Cue)를 제공하기 위해 사용자가 인식해야 하는 공의 회전 방향을 달리하는 방법을 사용하고 있지만, 회전 효과 외에도 여러 가지 방법을 추가로 CAPTCHA에 적용하여 자동화 공격 저항성을 높이기 위한 시도를 해볼 수 있다. Emerging Image 기법의 보안성은 인간의 인지 구조를 기계가 모방할 수 없다는 사실에 근거하고 있으며[15], 이는 기계학습 등의 컴퓨터 처리 기법을 통해 Emerging Image가 적용된 이미지에 포함된 요소에 대한 인식 혹은 판별이 가능하더라도 연산에 소요되는 자원 및 시간이 비효율적임을 의미한다. 또한, Emerging Image 효과 적용뿐만 아니라 DCG 요소를 추가한다면 단순히 CAPTCHA 화면의 객체를 인식하는 것만으로 CAPTCHA에 올바른 응답을 제출 할 수 없기 때문에 컴퓨터를 이용한 공격은 더더욱 힘들어지며, 설령 DCG 요소까지 감안한 자동화 공격을 수행하더라도 드래그 앤 드롭, 클릭 등의 동작들을 이미지 인식, 판별 기법과 조합하고 나서야 비로소 CAPTCHA 에 대한 응답이 생성되므로 자동화 공격에 대한 저항성을 추가로 확보 할 수 있다. 또한 가능한 모든 경우의 수를 입력해보는 단순한 자동화 공격에 저항하기 위한 방법으로 시도 횟수의 제한, 입력 행동 감지 등의 추가 보안 요소를 추가하는 것을 고려할 수 있다. 이는 새로운 CAPTCHA 메커니즘을 개발하는 것에 비해 비교적 낮은 비용으로 취할 수 있는 조치이므로, Emerging Image 기법이 적용된 CAPTCHA는 구현 방법에 따라 높은 자동화 공격 저항성을 가질 수 있음을 의미한다.

4.2 릴레이 공격 저항성

릴레이 공격 저항성 실험은 네트워크 대역폭과 네트워크 지연 시간이 통제된 환경에서 진행되었다. 피험자는 EIQ CAPTCHA가 실행되는 컴퓨터를 원격 조작하는 컴퓨터에서 EIQ CAPTCHA에 응답하였다. 해당 실험이 공격 실험이라는 사실은 데이터 편중 방지를 위해 사전에 설명하지 않았으며, 실험이 종료된 후에 실험의 목적을 설명하였다. 실험 과정에서 CAPTCHA 응답 소요시간, CAPTCHA 응답 성공, 실패 여부를 측정하였다.

대역폭 조절에는 NetLimiter4를 사용하여 대역폭을 제한하였으며, 네트워크 환경 시뮬레이션 프로그램인 clumsy 0.2를 사용하여 지연시간을 통제했다. 시뮬레이션 네트워크 대역폭은 중국의 평균 인터넷 대역폭인 3.7Mbps로 설정하였으며 네트워크 지연 시간은 중국의 평균 네트워크 지연 시간인 2885ms로 설정하였으나[17], RealVNC를 이용한 원격조작 환경에서 원격화면 갱신에 너무 긴 시간이 소요되

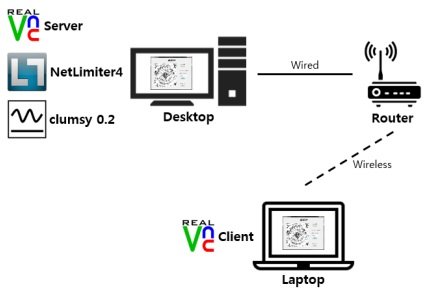


Fig. 6. Stream Relay Attack Environment

어 실험이 불가능하였다. 이에 실험이 가능한 수준인 1000ms(최초 설정 지연 시간의 약 35.71%)으로 지연시간을 낮추어 공격 실험을 진행하였다.

4.3 릴레이 공격 저항성 평가 결과

실험 결과, 실험 참여자 5명 모두 CAPTCHA응답에 실패하였으며, 룰렛 판에서 공이 멈춘 위치를 판별할 수 없어 Random guessing attack에 가까운 공격 양상을 보였다. 공격 실험에 소요된 시간과 결과는 아래 표와 같다. 이러한 결과는 움직임이 같은 3개의 공과 움직임이 다른 1개의 공이 같은 궤적 위에서 움직인 것과 낮은 네트워크 대역폭과 높은 네트워크 지연시간으로 인한 화면 끊김 현상이 원인으로 파악되었다.

Table 1. Stream Relay Attack Result

Person	Duration (ms)	Result
1	91199	Failure
2	95243	Failure
3	106477	Failure
4	103691	Failure
5	54240	Failure

V. 사용성 평가

5.1 평가 환경

EIQ CAPTCHA의 사용성 평가를 위해 기존의 Emerging Image Game CAPTCHA와 병행하여 사용해보는 평가 시나리오를 구성하였다. Emerging Image Game CAPTCHA와 EIQ CAPTCHA 양쪽을 모두 사용해본 후에는 피험자들에게 EIQ CAPTCHA에 대한 System Usability Scale(이하 SUS) 설문에 응답하도록 하였다. 평가 과정은 소음이 적은 강의실에서 실험 안내자와 피험자 1:1로 이루어졌으며, 실험 안내-설명-실험-설문 응답-종료 순서로 시행되었다. 실험 과정에서 수집한 정보는 각 CAPTCHA 응답에 소요된 시간, 응답 시도 횟수, 최종 성공 여부이다.

사용성 평가 실험은 피험자 내 설계 방식으로 구성하였다. 독립변수는 CAPTCHA의 형식(Scheme)이며, 종속변수는 CAPTCHA 응답 소요 시간과 오류율이다. 또한 학습 효과를 제거하기 위해 각 피험자에게 처음 사용하는 CAPTCHA의 종류에 대해 무작위 순서를 적용하여 평가를 진행하였다.

실험 진행 전, Emerging Image 효과와 각 CAPTCHA에 대한 간단한 사전 지식을 피험자에게 설명한 뒤 두 가지 CAPTCHA 모두에 5회씩 응답하도록 하였으며, 응답 과정에서 소요되는 시간, 시도횟수, 최종 결과를 파일에 로그로 남겼다. 평가 과정에서 EIQ 및 Emerging Image Game CAPTCHA 모두에 대해 피험자가 CAPTCHA 시작 버튼을 클릭했을 때부터 응답 시간을 측정하였다. 실험이 끝나면 10개의 항목으로 구성된 System Usability Scale(이하 SUS) 설문에 응답하도록 하였고 SUS 점수 역시 파일에 로그로 기록하였다.

5.2 평가 지표

실험 과정에서 측정한 요소는 주어진 CAPTCHA에 대해 응답하는데 소요된 시간, 응답 시도 횟수, 최종 성공 여부, SUS 점수이며, 실험 참여자가 CAPTCHA를 사용하는 과정에서 자동으로 로그 파일에 해당 지표가 기록되도록 EIQ CAPTCHA 프로토타입을 구현하였다.

5.3 평가 결과

실험에 참여한 총 인원 수는 17명으로, 남성 13명, 여성 4명으로 구성되었으며, 평균연령 29.29세, 표준편차는 8.47이다.

사용성 평가 결과, SUS 점수는 최대, 최솟값을 제거한 평균 57.67점, 표준편차 17.71로 측정되었다. 평균 응답 시간은 Emerging Image Game CAPTCHA가 약 17.82초, EIQ CAPTCHA가 15.61초를 기록하였다. 단, 실험 참여자들은 Emerging Image Game CAPTCHA에 대해 응

Table 2. Usability Test Result

EI Game Avg. Response Time (ms)	EIQ Avg. Response Time (ms)
17826.04	15612.40
EI Game Error rate (%)	EIQ Error rate (%)
89.35	9.41

답을 포기한 경우가 많았으며, 이는 Emerging Image Game CAPTCHA의 높은 Error rate(89.35%)로 설명할 수 있다. 반면, EIQ CAPTCHA의 Error rate는 9.41%로, Emerging Image Game CAPTCHA에 비해 획기적으로 낮음을 알 수 있다. 또한, 본 실험에서 측정한 Error rate는 시도 횟수를 1회로 제한했을 경우에 측정된 값이며, 시도 횟수에 제한을 두지 않을 경우 EIQ CAPTCHA의 Error rate는 0%로 모든 사용자가 최종적으로 응답에 성공하는 양상을 보였다.

CAPTCHA의 종류에 따른 응답 시간의 통계적 유의성을 검증하기 위하여 대응표본 t 검정을 수행하였으나 응답 시간에 대한 유의성 검증에 실패하였다. 이는 Emerging Image Game CAPTCHA의 응답 시간 값이 사용자가 CAPTCHA에 올바른 응답을 할 때 까지 소요된 시간이 아니기 때문에 발생한 문제로, 실제로 측정된 응답 시간들 중 대다수는 사용자가 Emerging Image Game CAPTCHA 응답을 포기할 때 까지 소요된 시간이기 때문이다.

Error rate에 대한 대응표본 t 검정 결과, CAPTCHA의 종류에 따라 Error rate에 유의한 차이가 있음이 밝혀졌다. ($t(16) = -15.033$, $p<0.001$)

VI. 결 론

본 연구에서는 기존에 존재하던 CAPTCHA의 종류와 여러 가지 CAPTCHA들에 대한 공격 사례에서 공격 방식, 성공률에 대해 알아보았다. 또한 이러한 선행 연구를 통해 현존하는 CAPTCHA들이 각종 자동화 공격, 정적 릴레이 공격, 스트림 릴레이 공격에 취약할 수 있음을 보이고, 이러한 공격에 대해 저항하기 위해 제안된 Emerging Image Game CAPTCHA에 대해 중점적으로 살펴보았다. Emerging Image Game CAPTCHA는 높은 공격 저항성을 보인 대신, 사용성이 나쁘다는 한계점을 갖고 있었다. 본 연구에선 사용성을 개선한 EIQ CAPTCHA와 EIQ CAPTCHA의 설계 방법을 제안하고 제안한 CAPTCHA에 대한 사용성 평가와 릴레이 공격 저항성을 Lab test를 통해 검증하였다.

본 연구의 시사점으로는 기존에 제시되었던 Emerging Image 기반 CAPTCHA 보다 사용성이 개선된 EIQ CAPTCHA와 사용성을 저해하지 않는 EIQ CAPTCHA 설계 방법을 제안했다는 데에 그 의미를 둘 수 있다. 이 설계 방법은 Emerging Image 기법을 CAPTCHA 구성 요소 중 사용자가 판독해야 하는 요소에 직접적으로 적용하지 않고, 판독해야 하는 요소에 대한 힌트 혹은 실마리를 제공할 수 있도록 적용하는 방법이다.

또한, Stream Relay 공격 Lab test를 통해 EIQ CAPTCHA는 Stream Relay 공격에 대한 저항성을 가진다는 것을 확인할 수 있었다. CAPTCHA 별 평균 응답 시간 차이는 Emerging Image Game CAPTCHA의 저조한 정답률로 인해 유의한 결과를 도출할 수 없었으나, EIQ CAPTCHA는 Emerging Image Game CAPTCHA 보다 획기적으로 낮은 Error rate를 보였고, 이러한 경향은 통계적으로 유의함을 확인할 수 있었다.

한계점으로는 3장 후반부에 서술했듯이, 본 연구에서 제안하는 EIQ CAPTCHA와의 비교 대상인 Emerging Image Game CAPTCHA의 원본을 입수하지 못했기 때문에 사용성 평가의 신뢰성에 한계가 있다. 본 연구에서 사용성 평가를 진행하기 위해 직접 개발하여 실험 과정에 반영한 Emerging Image Game CAPTCHA는 원본의 작동 방식에 비해 주요 메커니즘의 대부분이 생략 및 간략화 되어 있기 때문이다. 이러한 문제점은 CAPTCHA 원본을 실험에 반영함으로써 극복할 수 있을 것이다.

릴레이 공격 저항성 실험에서 사용된 시뮬레이션 변수들은 중국의 주요 네트워크 지표들로, 다양한 네트워크 환경들에 대한 실험이 이루어지지 못했다. 이러한 한계점을 보완하려면 제 3자 CAPTCHA 해결 서비스를 하고 있는 미국, 인도 등의 다양한 국가의 네트워크 환경을 반영할 수 있는 시뮬레이션 변수에 대한 검증이 필요하다. 또한 실제로 서비스되고 있는 제3자 CAPTCHA 해결 서비스들에 대한 기반 구조, 알고리즘, 메커니즘 등에 대한 분석이 뒤따라야 보다 완성된 릴레이 공격 저항 연구가 될 수 있을 것이다.

본 연구에서 진행한 사용성 평가는 Lab test 로 진행되었으며, 이 과정에서 모집단의 표준화가 충분히 고려되지 못했다는 한계점을 갖는다. 또한 실험 참여자의 수는 17명으로, 실험 결과에 대한 신뢰성이 보장되기 어려운 숫자이기에 차후 진행될 연구에서는 이러한 부분을 보완 및 고려하여 추가 실험을

진행해야 할 필요성이 있다.

사용성 평가 결과, SUS 점수는 평균 57.67점으로, 일반적인 사용성 기준인 68점보다 낮아 사용하기 다소 불편하다는 결과가 도출되었다. 보다 개선된 사용성을 위해 룰렛만이 아닌, 다른 방식의 게임들에 대해서도 사용성 테스트를 해볼 필요성이 있으며, 앞서 언급한 Emerging Image Game CAPTCHA의 원본을 확보하여 DCG 요소를 추가한 사용성 테스트 또한 필요하다.

결론적으로 EIQ CAPTCHA는 기존에 제안되었던 Emerging Image Game CAPTCHA보다 개선된 사용성, 낮은 Error rate를 보였다. 본 연구의 사용성 평가 및 실험들은 Lab test로 진행되어 모집단의 수가 적고 표준화에 대한 고려가 되지 않았다는 한계점을 갖지만, 이로 인해 얻은 결과는 추후에 진행될 추가 실험에 대한 밑거름이 될 수 있으며, 추후 진행될 연구에서 DCG를 도입한 EIQ CAPTCHA와 완전한 Emerging Image Game CAPTCHA에 대한 사용성 평가가 후속되고, 원격 공격 시뮬레이션의 인자 다양화, 다양한 Emerging Image 파라미터 적용 등의 다양한 실험을 추가로 진행함으로써 본 연구의 미비한 부분을 보완할 수 있을 것이다.

References

[1] Gao, S., Mohamed, M., Saxena, N., & Zhang, C. "Emerging image game CAPTCHAs for resisting automated and human-solver relay attacks." In Proceedings of the 31st Annual Computer Security Applications Conference, ACM, pp. 11-20, Dec. 2015.

[2] Von Ahn, L., Blum, M., Hopper, N. J., & Langford, J. "CAPTCHA: Using hard AI problems for security," In International Conference on the Theory and Applications of Cryptographic Techniques, Springer Berlin Heidelberg, pp. 294-311, May. 2003.

[3] Mori, G., & Malik, J. "Recognizing objects in adversarial clutter: Breaking a visual CAPTCHA," In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference, IEEE. Vol. 1, pp. I-I, Jun. 2003.

[4] Moy, G., Jones, N., Harkless, C., & Potter, R. "Distortion estimation techniques in solving visual CAPTCHAs," In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on Vol. 2, IEEE, pp. II-II, Jun. 2004.

[5] Bansal, Abhay, et al. "BREAKING A VISUAL CAPTCHA: A NOVEL APPROACH USING HMM."

[6] Yan, J., & El Ahmad, A. S. "A Low-cost Attack on a Microsoft CAPTCH," In Proceedings of the 15th ACM conference on Computer and communications security, ACM. pp. 543-554, Oct. 2008.

[7] Chellapilla, K., & Simard, P. Y. "Using machine learning to break visual human interaction proofs (HIPs)," In NIPS, pp. 265-272, Dec. 2004.

[8] Xu, Y., Reynaga, G., Chiasson, S., Frahm, J. M., Monrose, F., & Van Oorschot, P. C. "Security analysis and related usability of motion-based captchas: Decoding codewords in motion," IEEE transactions on dependable and secure computing, 11(5), pp. 480-493, Sep. 2014.

[9] Elson, J., Douceur, J. R., Howell, J., & Saul, J. "Asirra: a CAPTCHA that exploits interest-aligned manual image categorization," In ACM Conference on Computer and Communications Security, Vol. 7, pp. 366-374, Oct. 2007.

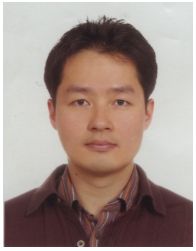
[10] Golle, P. "Machine learning attacks against the Asirra CAPTCHA," In Proceedings of the 15th ACM conference on Computer and communications security, ACM, pp. 535-542, Oct. 2008.

[11] Qiujie, L., Yaobin, M., & Zhiquan, W. "A survey of CAPTCHA technology," Journal of Computer Research and Development, Vol.49, no.3, pp. 469-480, Mar. 2012.

〔12〕 Ross, S. A., Halderman, J. A., & Finkelstein, A. "Sketcha: a CAPTCHA based on Line Drawings of 3D Models," In Proceedings of the 19th international conference on World wide web, ACM, pp. 821-830, Apr. 2010.
〔13〕 Nguyen, V. D., Chow, Y. W., & Susilo, W. "On the security of text-based 3D CAPTCHAs," Computers & Security, Vol.45, pp. 84-99, Sep. 2014.
〔14〕 Mohamed, M., Sachdeva, N., Georgescu, M., Gao, S., Saxena, N., Zhang, C., ... & Chen, W. B. "Three-way dissection of a game-captcha: Automated attacks, relay attacks, and usability," arXiv preprint arXiv:1310.1540, Oct. 2013.
〔15〕 Mitra, N. J., Chu, H. K., Lee, T. Y., Wolf, L., Yeshurun, H., & Cohen-Or, D. "Emerging images," In ACM Transactions on Graphics (TOG), ACM, Vol. 28, No. 5, pp. 163, Dec. 2009.
〔16〕 Yang, C. H., Kuo, Y. M., & Chu, H. K. "Synthesizing Emerging Images from Photographs," In Proceedings of the 2016 ACM on Multimedia Conference, ACM, pp. 660-664, Oct. 2016.
〔17〕 Quarter, Akamai Releases Third. "State of the Internet Report." 2015.

〈저자소개〉

양 원 석 (Wonseok Yang) 학생회원
2015년 2월: 상명대학교 미디어소프트웨어학과 학사
2015년 3월~2017년 2월: 연세대학교 정보대학원 석사과정
〈관심분야〉 Mobile Security, Usable Security 등



권 태 경 (Taekyoung Kwon) 종신회원
1992년 2월: 연세대학교 컴퓨터과학과 학사
1995년 2월: 연세대학교 컴퓨터과학과 석사
1999년 8월: 연세대학교 컴퓨터과학과 박사
1999년~2000년: U.C. Berkely Post-Doc.
2001년~2013년 8월: 세종대학교 컴퓨터공학과 교수
2007년~2008년: Univ. Maryland at College Park 교환교수
2013년 9월~현재: 연세대학교 정보대학원 교수
〈관심분야〉 암호프로토콜, 네트워크 프로토콜, 사물인터넷 보안, HCI 보안 등