Journal of The Korea Institute of Information Security & Cryptology(JKIISC) ISSN 1598-3986(Print)
VOL.23, NO.4, August 2013 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2013.23.4.571

콘텐츠 기반 캡차를 이용한 인터넷 뱅킹 서비스의 보안성 향상 기법*

이 상 호,[1†] 김 성 호,[1] 강 전 일,[1] 변 제 성,[1] 양 대 헌,[1] 이 경 희[2‡]
[1]인하대학교, [2]수원대학교

A Method of Enhancing Security of Internet Banking Service using Contents-Based CAPTCHA*

Sang-ho Lee,[1†] Sung-ho Kim,[1] Jeon-il Kang,[1] Je-sung Byun,[1] Dea-hun Nyang,[1]
Kyung-hee Lee[2‡]
[1]INHA University, [2]The University of Suwon

요 약

인터넷 뱅킹은 시간에 얽매이지 않는 편리성 때문에 널리 사용되고 있다. 하지만 인터넷 뱅킹 서비스가 자동화 프로그램에 의해 공격이 가능해 진다면 수많은 계좌가 공격의 위험에 노출될 것이며, 그로 인한 피해는 천문학적인 금액이 될 것으로 예상된다. 이와 같은 공격에 대응하기 위하여 Arcot사와 MS사는 각각 VPS와 MS 워터마크를 고안하였고, 이는 계좌 이체 정보를 이용하여 문자열 기반 캡차를 생성하고 있다. 이 논문에서는 위 기술들이 국내 인터넷 뱅킹 서비스에 적용되었을 시 발생할 수 있는 취약점을 살펴보고, 기존의 기술을 개선시킨 대응방법과 사용자와 서버 사이에 알려진 계좌 이체 정보를 이용하는 콘텐츠 기반 캡차 생성을 통해 인터넷 뱅킹의 안전성을 강화하는 기법을 제안한다.

ABSTRACT

Internet banking service has a advantage that is unrestricted by time. If automated programs are able to attack Internet banking services, a number of accounts can be attacked at the same time and as a result, damage will be considerably increased. To cope with such attacks, two methods, VPS and MS watermark, were introduced by Arcot and MS respectively. The methods use text-based CAPTCHAs in the process of transfer approval to distinguish automated programs from legal human users. In this paper, we point out the security threats of the methods when those are applied to Internet banking services. Especially, we consider the attack that are performed by extract specific string from text-based CAPTCHAs and it's countermeasure. Also we suggest a method of enhancing security of internet banking services. Our method is based on contents-based CAPTCHAs that are consist of known transfer information between user and server.

Keywords: CAPTCHA, Internet banking, MITB, secure card, OTP

접수일(2013년 1월 14일), 수정일(1차: 2013년 3월 27일, 2차: 2013년 5월 20일), 게재확정일(2013년 6월 13일)
* 이 논문은 인하대학교의 지원에 의하여 연구되었음.

† 주저자, 181cm76kg245@gmail.com
‡ 교신저자, khlee@suwon.ac.kr(Corresponding author)

I. 서 론

인터넷 뱅킹 서비스는 금전적인 부분과 연관되기 때문에 무엇보다 보안이 필수적이다. 인터넷 뱅킹 서비스를 제공하는 업체들은 잠재적인 위협에 대응하기 위해 보안 프로그램을 사용자에게 설치하고 사용할 것을 강요하는 방식으로 자사의 인터넷 뱅킹 서비스를 보호하고 있다. 하지만 제공되는 보안 프로그램들이 악성 프로그램들에 의해 변조되어 비정상적인 작업을 실행할 경우, 보안카드와 OTP 같은 추가적인 보안 매체를 사용한다고 하더라도 인터넷 뱅킹 서비스 이용 시 발생되는 트랜잭션에 대해서는 무결성 검증을 수행하지 않기 때문에 트랜잭션을 변조하는 공격에 매우 취약하다. 이와 같은 취약점을 이용하여 악성코드에 감염된 보안 프로그램이 트랜잭션을 변조하는 공격을 수행한다면 그 피해는 사용자에게 돌아갈 것이 분명하다. 더불어 이러한 공격이 자동화 되어 실행된다면 공격의 범위는 더 넓어질 것이며, 피해 금액 또한 공격자가 직접 개입하는 방식의 공격보다도 더 클 것이다.

일반적으로 캡차(CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart)는 컴퓨터와 사람을 식별할 수 있게 해주는 공개된 튜링 테스트를 의미한다[1]. 이러한 캡차의 특징을 이용하여 악성 프로그램에 의한 거래 변조를 사용자가 인식 하고 비정상적인 계좌 이체가 수행되는 것을 사전에 피할 수 있을 것이다. 대표적으로 Arcot사의 VPS[2]와 MS사의 워터마크[3]가 있고, 이들 모두 프로그램이 문자열 캡차를 인식하지 못한다는 점에 초점을 맞추고 있다. 하지만 VPS와 MS 워터마크의 승인 페이지에 사용된 캡차에서 특정 문자열을 추출해 낼 수 있다면 이체 승인 페이지 또한 악성 프로그램에 의도대로 변조될 수 있는 가능성은 남아 있다.

이 논문에서는 위와 같은 트랜잭션을 변경하는 자동화된 공격에 대응하기 위해 기존의 계좌이체 승인페이지에 캡차를 적용하는 기법으로서 기존의 계좌이체가 가지는 취약점을 보안하고자 한다.

이 논문의 구성은 다음과 같다. 2장에서는 인터넷 뱅킹 서비스의 위협 사례를 살펴보고 위협 모델을 정의한다. 3장에서는 캡차를 이용하여 인터넷 뱅킹 서비스의 보안성을 강화한 기존의 기술에 대해 살펴보고, 4장에서는 기존의 기술이 가지는 문제점을 보안한 콘텐츠 기반 승인 페이지 구성 방법을 제시한다. 5장을 통해 이 논문의 결론과 향후연구를 싣는다.

II. 인터넷 뱅킹 서비스의 위협 모델

이 장에서는 인터넷 뱅킹 서비스의 공격 사례에 대해서 살펴보고 이 논문에서 고려하고 있는 위협 모델을 정의한다.

2.1 해외 사례

2009년 웹 보안 업체인 Finjan사에서는 인터넷 뱅킹 해킹 사건에 대한 기술문서[4]를 공개하였다. 이 문서에 따르면 공격자는 웹브라우저의 보안 취약점을 이용하여 사용자들의 단말기에 악성 프로그램을 설치한 후, 봇넷 형태로 단말기들을 관리했다. 대략 6,400개의 단말기에 악성 프로그램이 설치되었고, 그 중 수백 대의 단말기에서 악의적인 계좌 이체가 발생하였으며, 22일 동안 악성 프로그램에 의해 발생한 피해 금액은 대략 300,000 유로에 이른다고 밝혔다. 위 사례에서의 특징은 다음과 같다.

- 공격자는 OTP 및 SSL/PKI 등을 우회하기 위해서 MITB(Man-in-the-Browser)[5] 공격을 하였고, 계좌 이체 시 사용자에 의해 발생되는 파라미터(수신 계좌, 계좌이체 금액)를 변조하여 악의적인 계좌이체를 수행했다.
- 감염된 다수의 단말기들을 봇넷 형태로 관리했고, 악성 프로그램은 명령 서버를 통해 이체 액수와 관련된 정보 및 송신 계좌 등을 전달받았다. 악의적인 이체가 수행되는 과정은 공격자의 개입 없이 자동화로 진행되었다.
- 온라인 사기 방지 시스템(anti-fraud)에 탐지되는 것을 최소화하기 위해 다음의 두 가지 사항을 고려하여 계좌 이체를 수행 하였다. 이체 수행 전후에 계좌 잔고가 남아 있는지를 확인하고, 잔고가 남아 있도록 계좌이체를 수행한다. 계좌이체 액수를 높게 책정하지 않고, 매번 계좌이체 금액을 특정 범위 내에서 무작위로 선택한다.
- 악의적인 계좌 이체를 숨기기 위해 계좌이체 조회와 같은 인터넷 뱅킹 서비스 페이지 역시 조작하여 사용자에게 보여주었다. 이를 지속하기 위해서 사용자 단말기에 이와 관련된 설정 파일을 유지했다.

2.2 국내 인터넷 뱅킹 서비스 위협 사례

2010년 맹영재 등은 2.1절에서 설명한 MITB 공격이 실제 국내 인터넷 뱅킹 환경에서 실행될 수 있음을 증명하였다[6]. 이 논문 또한 계좌 이체 시에 발생되는 트랜잭션에 대한 검증이 수행되지 않는 점을 이용 하였으며 웹 보안, 개인 방화벽, 키보드 해킹방지 및 암호화 프로그램이 설치되고 실행 중 이더라도 이러한 공격이 가능하다는 점을 보였다.

맹영재 등은 BHO(Brower Helper Object)를 이용한 자동화 악성코드를 제작하여 화면을 변조하는 방법으로 공격을 시도하였다. BHO는 계좌 이체 페이지와 똑같은 입력 폼을 웹브라우저 최상단에 노출시키고 이 노출된 폼에 사용자로부터 이체계좌번호와 이체금액을 입력받고 별도의 공간에 저장하였으며, 계좌이체 비밀번호와 같은 사용자만 알고 있는 비밀은 서버에서 제공한 폼에 입력되도록 한 후 별도의 조작을 가하지 않고 서버에 그대로 전송하였다. 사용자가 입력한 내용을 서버에 전송하기 전에 악성코드는 이체계좌번호 입력란에 BHO를 통해 사전에 정의된 값을 입력하였으며, 이체금액은 은행 사에서 제공하는 함수를 이용하여 이체가능 전액으로 수정하였다. 계좌 이체 내역 페이지로 이동하게 되면 별도의 공간에 저장되어 있던 사용자 입력 정보를 화면에 덮어쓰는 방식으로 사용자의 눈을 속였다.

이 공격은 모든 보안 프로그램을 우회하여 자동으로 공격이 가능하였으며, 보안 카드나 OTP등을 사용하더라도 인터넷 뱅킹이 안전하지 않다는 점을 증명하였다.

2.3 인터넷 뱅킹 서비스의 위협 모델

이 논문에서는 가정하고 있는 위협 모델은 다음과 같다.

- 일련의 공격 과정들은 악성 프로그램에 의해 자동으로 수행된다.
  인터넷 뱅킹 서비스에 이용되는 단말기를 공격자가 직접 개입하여 시도되는 공격의 경우, 동시에 공격할 수 있는 대상의 수가 한정적일 수밖에 없다. 하지만 공격이 프로그램에 의해 자동화될 경우에는 수많은 단말기가 동시에 공격 받을 수 있으며, 그로 인해 발생되는 피해 규모는 공격자가 직접 개입하는 공격보다 더 클 것이다. 이 논문에서는 후자의 경우에 대해서 고려한다.
- 악성 프로그램은 인터넷 뱅킹 서비스를 보호하는 보안 프로그램을 우회할 수 있다.
  악성 프로그램이 단말기를 장악했다는 의미는 보안 프로그램 또한 악성 프로그램에 의해 변조될 수 있다는 것을 의미한다. 변조를 통해 악성 프로그램은 자신의 존재를 숨기거나 특정 프로그램들의 기능을 제한할 수 있다.
- 공격자는 사용자만 알고 있는 정보(송·수신 계좌, 비밀번호, 계좌이체 금액 등)와 더불어 암호화, 복호화에 사용되는 비밀 키를 포함한 정보 또한 수집 및 변조가 가능하다.

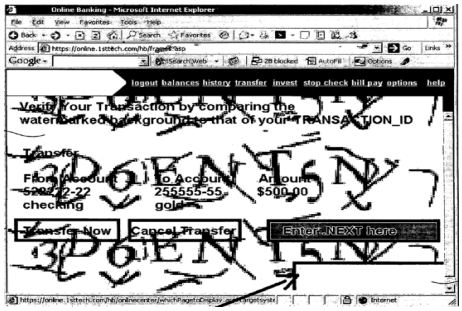
2.4 위협 모델의 성공적인 공격 방법

일반적으로 사용하는 보안카드의 경우 35개의 번호가 반복적으로 사용되기 때문에 임의의 단말기가 악의적인 목적으로 공격자에 의해 장악되었을 상황을 가정하면 공격자는 사용자가 입력하는 정보를 관찰하여 보안카드의 모든 정보를 획득할 수 있고, 공격자 스스로 악의적인 계좌 이체를 수행할 수 있다. 단, OTP는 100만개 가량의 번호가 사용되기 때문에 OTP 생성기를 물리적으로 획득하지 않는 한 계좌 이체 시에 필요한 OTP 값을 알아내기 힘들다. 이와 같은 경우, 공격자는 스스로 계좌 이체를 수행할 수는 없지만 사용자가 계좌 이체 수행 시 트랜잭션에 대한 검증을 수행하지 않는 다는 점을 이용하여 트랜잭션 변경을 통한 악의적인 계좌 이체를 수행할 수 있다.

OTP는 생성 방식에 따라 비동기화 방식과 동기화 방식으로 나눌 수 있다. 비동기화 방식은 사용자 정보로 부터 OTP 값을 생성하는 방식으로 계좌이체와 관련된 정보를 사용할 경우 이체 정보와 연관된 OTP 값을 생성해 낼 수 있다. 동기화 방식의 경우에는 OTP 기기와 인증 서버 간에 동기화된 정보(시간 정보, 카운터 값)를 이용하여 OTP 값을 생성하기 때문에 계좌 이체 정보와 같은 특정 정보와 연관된 OTP 값을 생성하는 것은 사실상 불가능하다.

계좌 이체 승인 시에 사용되는 OTP 값이 계좌 이체 정보와 연관되어 생성된 값이 아닌 OTP 기기의 소유의 여부를 확인하기 위해서만 사용되고 계좌 이체 시에 발생되는 트랜잭션에 대한 검증이 이루어지지 않는다면 이 역시 MITB 공격에 노출될 수 있다.

(a) VPS의 이체 승인 페이지



(b) MS 워터마크의 이체 승인 페이지

〔그림 1〕 이체 승인페이지에 캡차를 적용한 기존 기술

III. 관련 연구 및 동향

인터넷 뱅킹 서비스를 보호하기 위한 기술로는 Arcot사의 VPS와 MS사의 MS 워터마크가 있다. 이 장에서는 VPS와 MS 워터마크에 대해 살펴보고, 이 기술들의 취약점에 대해 살펴본다.

3.1 캡차를 이용하여 승인 페이지를 구성한 기존 기술

VPS와 MS 워터마크는 문자열 왜곡, 노이즈, 배경 그라데이션 등을 추가하여 자동화 프로그램이 캡차 문자열을 인식하는 것을 어렵게 하고 있다. 더불어 사용자에게 보이는 페이지에 무작위성이 존재 하지 않는다면 이 또한 쉽게 변조 될 수 있다. 따라서 사용자는 인식할 수 있지만 일괄 되지 않는 무작위성이 있어야 할 것이다. VPS와 MS 워터마크는 이러한 인식 가능한 무작위성 캡차을 제공하고 있다.

3.1.1 VPS의 승인 페이지 및 계좌 이체 승인 과정

VPS에서는 〔그림 1-a〕와 같이 계좌 이체 정보 및 서버에서 생성한 OTP 문자열을 기반으로 캡차를 생성한 후, 이를 승인페이지에 추가하여 사용자에게 전송한다. 사용자는 캡차에 적힌 계좌 이체 정보를 확인한 후 이체 정보가 정상적이라고 판단되면 캡차에 적힌 OTP 문자열을 기입하여 서버로 전송한다. 이후 서버는 사용자가 보낸 문자열과 OTP를 비교하여 두 문자열이 일치한다면 계좌 이체를 수행한다.

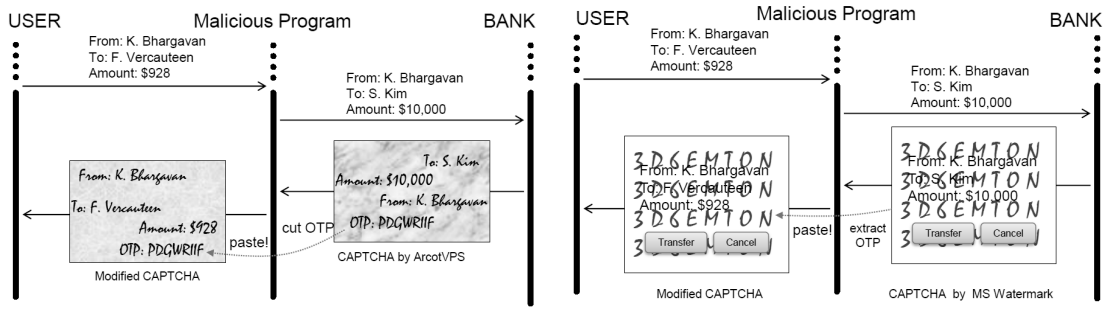
3.1.2 MS 워터마크의 승인 페이지 및 계좌 이체 승인 과정

MS 워터마크에서는 〔그림 1-b〕와 같이 이체 승인 페이지 위에 OTP 문자열을 덮어 씌워 사용자에게 전송한다. 사용자는 서버와 동기화된 OTP 생성기와 같은 보안 장비를 항상 소지하고 있어야 한다. 서버로부터 승인 페이지를 전송받은 사용자는 계좌 이체 정보가 정상인지 확인함과 동시에 덮어 씌워진 OTP 문자열과 보안 장비에 표시된 문자열이 동일한지를 확인한다. 확인 후 사용자는 보안 장비를 통해 새로 생성된 OTP 문자열을 입력하여 서버로 전송한다. 사용자가 전송한 OTP 문자열 즉, 보안장비를 통해 새로 생성된 문자열과 서버에서 새로 생성된 OTP 문자열이 일치한다면, 서버는 최종적으로 계좌 이체를 수행한다.

특허 출원 중인 Arcot의 VPS와 MS 워터마크는 브라우저에서 BHO(Brower Helper Object)로 인한 모든 트랜잭션 변경 사항을 사용자에게 보여준다. Arcot VPS는 트랜잭션에 대한 정보와 트랜잭션 확인 코드를 동일한 웹 채널을 통해 캡차 형식으로 사용자에게 제시하고, 사용자는 캡차를 읽고 트랜잭션 정보를 확인한 후 정해진 시간 내에 확인 코드를 입력하여 트랜잭션을 진행해야 한다. 일반적으로 MITB 트로이 목마는 정해진 시간 내에 캡차를 구문 분석하고 트랜잭션 확인 코드를 읽을 수 없다. MITB로 인해 트랜잭션이 수정된 경우 수정된 트랜잭션 전체가 캡차 이미지로 표시되고 따라서 사용자는 원래 입력한 정보와 비교하여 변경된 내용을 쉽게 찾아 트랜잭션을 중지 할 수 있다. 이는 맹영재 등이 BHO(Brower Helper Object)를 이용한 자동화 악성코드를 제작하여 화면을 변조하는 방법과 같은 MITB 공격을 막을 수 있다.

3.2 기존 기술들의 보안성 고찰

VPS와 MS 워터마크에서는 악성 프로그램이 캡차에 쓰인 문자열을 정확하게 읽어내지 못하는 점을 이



(a) VPS의 이체 승인 페이지 변조 (b) MS 워터마크의 이체 승인 페이지 변조

〔그림 2〕 악성 프로그램이 캡차를 변조하는 방법. 악성 프로그램이 캡차 문자열을 읽을 수 없다 하더라도 OTP를 추출할 수 만 있다면 사용자를 속일 수 있다.

용하여 승인 페이지의 보안성을 향상시켰다. 또한 문자열에 다양한 필터(문자열 왜곡, 노이즈 추가, 문자 포갬, 동일 색상 문자, 배경 그라데이션 등)를 적용하여 OTP 문자열을 읽어내는 것을 더욱 어렵게 만들었다.

VPS와 MS 워터마크는 이체 정보와 OTP를 캡차로 만들거나 보안장비로 확인 가능한 OTP를 캡차로 만들어 전송함으로써, 앞서 언급하였던 승인 페이지의 변조 방지를 위한 필요조건인 '인식 가능한 무작위성'을 만족하였다. 그러나 이는 단지 '필요조건'일 뿐이며 승인 페이지가 변조 되는 것을 막았다고는 볼 수 없다. OTP를 이용한 캡차만으로 승인 페이지가 변조되는 것을 막기 위한 '충분조건'이 되기 위해서는 악성 프로그램의 의도대로 캡차를 변조하는 것이 불가능해야 한다. 기존의 방법들은 OTP의 유효성만 검사하기 때문에 캡차에서 OTP만을 성공적으로 분리해낼 수 있다면 악성 프로그램은 승인 페이지를 의도대로 변조할 수 있다.

악성 프로그램의 의도대로 캡차를 변조하기 위해서는 캡차 문자열로부터 OTP를 추출해 내는 작업이 필요하다. 〔그림 2-b〕의 MS 워터마크의 경우, 이체 정보는 인쇄체 문자의 형태를 띠고 있으며, OTP 문자열은 악성 프로그램에 의해 읽혀지지 않기 위해 각종 필터들이 적용되어 있지만 페이지의 특정 위치에 반복적으로 나타나고 있다. 이러한 특징을 이용해 다음과 같이 OTP 문자열을 추출할 수 있다.

- 색상 정보를 이용한 OTP 추출
  MS 워터마크의 경우 OTP 문자열 색상이 계좌 이체 정보위에 덮여 써졌기 때문에, 서로 글자 색상을 다르게 하는 방식을 이용하여 사용자 편의성을 향상시켰다. 하지만 OTP와 배경 색상이 서로 다른 색상으로 구분되어 있을 경우 픽셀의 색상 정보를 이용하여 특정 색상(계좌 이체 정보를 나타내고 있는 검은색 픽셀을 제외한 픽셀)을 추출하거나 군집화 알고리즘을 이용하여 픽셀을 두 군집으로 분할한 후, 군집을 이루는 픽셀의 수가 많은 군집을 선택하여 OTP를 추출할 수 있다.
- 반복되는 위치 정보를 이용한 OTP 추출 (색상 정보가 없을 경우)
  MS 워터마크에서는 계좌 이체 정보는 인쇄체 문자의 형태로 특정 위치에 나타나고 있고 OTP 문자열은 반복적으로 사용되며 적용된 노이즈 또한 동일하게 사용되고 있다. 문자열들을 구성하는 픽셀들의 위치에 무작위성이 존재하지 않고 특정 패턴이 나타나기 때문에 이런 정보를 이용하여 OTP를 추출할 수 있다.

악성 프로그램이 이와 같은 특징을 이용하여 OTP를 추출하고 난 이후에는, OTP와 사용자가 입력한 계좌 이체 정보로 새로운 캡차를 생성하여 사용자의 눈을 속일 수 있다.

〔그림 2-a〕의 VPS의 경우 계좌 이체 정보 와 OTP 문자열에 각각 일정하지 않은 폰트와 필터가 적용되어 있다. 또한 OTP 문자열의 위치가 고정되어 있지 않기 때문에 MS 워터마크에서와 같이 OTP 값을 추출해내기는 어려워 보인다. 하지만 이체 정보의 각각의 요소들(수신자, 송신자, 이체 금액, OTP)이 일정한 거리를 두고 있고, 이러한 특징을 이용하여 각 요소들을 분리한다면 OTP 문자열을 추출하는 것은 분리된 개별 요소들 중에서 OTP를 선택하는 문제로 바뀌게 된다. 공격자 입장에서는 이체 요소의 문자열 길이와 같은 부가적인 정보들을 알고 있다면 이를 이용하여 OTP를 선택할 수 있다.

IV. 콘텐츠 기반 캡차를 이용한 승인 페이지 보호

자동화 악성 프로그램이 캡차를 정확하게 읽지 못하거나 읽을 확률이 낮다는 점 때문에 캡차가 사용되고 있지만 많은 연구를 통해 캡차가 읽혀질 수 있음은 이미 증명되었다[7]. 이와 같이 캡차를 읽는 공격에 대응하기 위해서 필터의 강도를 높여 사용할 수 있겠지만, 사람조차 읽기 어려운 수준까지 도달할 경우 캡차의 본래 기능이 저하되기 때문에 필터의 적용이 제한적일 수밖에 없다. 하지만 사용자와 검증하는 매체 사이에 서로 공유되는 정보를 캡차로 생성한다면 필터의 적용 강도가 제한되는 문제를 극복할 수 있다. Harada 등의 논문[8]에서는 불분명해 보이는 이미지가 제시되었을 때 원본 이미지에 대한 사전 정보가 없는 사용자는 이를 유추할 수 없지만, 원본 이미지에 대한 정보를 가지고 있는 사용자는 불문명한 이미지가 무엇인지 판단한 수 있다고 밝혔다.

4.1 사용자 정보 유무에 따른 캡차 인식률 실험 - 보다 어려운 문자열 캡차의 도입

4.1.1 정보의 양에 따른 캡차 인식률

Harada 등이 논문에서 밝힌 내용을 캡차에 적용하여도 같은 결론을 낼 수 있는지를 확인하기 위해 15명을 대상으로 실험을 진행 하였다. 실험은 영문, 숫자, 또는 영문과 숫자 4자리로 이루어진 서로 다른 기존의 캡차에 동일한 왜곡을 심하게 주어 [그림 3]과 같이 생성하고 피 실험자에게 캡차에 대한 아무런 정보도 주어지지 않은 경우([그림 3]과 같은 캡차를 피 실험자에게 제공한 후 캡차의 문자열을 인식 하는지 관찰), 캡차에 대한 약간의 정보를 제공한 경우([그림 3]과 같은 캡차를 제공하고 예를 들어 알파벳 d로 시작하는 캡차를 선택하라는 지문을 제공한 후 캡차의 문자열을 인식하는지 관찰), 그리고 캡차에 대한 모든 정보를 제공한 경우([그림 3]과 같은 캡차를 제공하고 예를 들어 캡차의 문자열이 9ki7인 캡차를 선택하라는 지문을 제공한 후 캡차의 문자열을 인식하는지 관찰) 각각에 대해 5개의 보기를 제공하고 인식률과 시간을 측정하였다.

[그림 3] 문자열에 왜곡을 심하게 주어 생성한 캡차

실험의 결과를 [표 1]에 정리하였다.

[표 1] 정보의 양에 따른 캡차 인식률

(a) 정보의 양에 따른 캡차 인식 실험

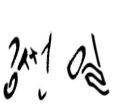
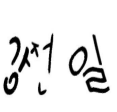
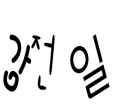
	실험 1 (정보 제공 없음)	실험 2 (약간의 정보를 제공)	실험 3 (정보 모두 제공)
관찰 1	2 / 5	1 / 5	3 / 5
관찰 2	0 / 5	0 / 5	1 / 5
관찰 3	0 / 5	2 / 5	4 / 5
관찰 4	1 / 5	3 / 5	5 / 5
관찰 5	0 / 5	4 / 5	5 / 5
관찰 6	0 / 5	4 / 5	5 / 5
관찰 7	1 / 5	2 / 5	4 / 5
관찰 8	0 / 5	3 / 5	4 / 5
관찰 9	0 / 5	3 / 5	3 / 5
관찰 10	0 / 5	3 / 5	5 / 5
관찰 11	0 / 5	2 / 5	4 / 5
관찰 12	0 / 5	1 / 5	4 / 5
관찰 13	0 / 5	3 / 5	5 / 5
관찰 14	0 / 5	3 / 5	4 / 5
관찰 15	0 / 5	3 / 5	4 / 5

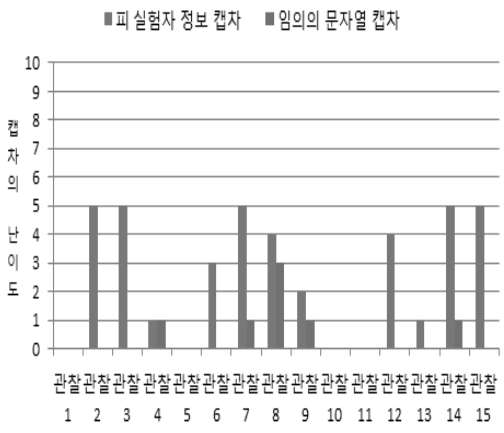
(b) 정보의 양에 따른 캡차 인식 실험 결과

	인식률 (success rate)	평균 시간(average time per round)
실험 1	4/75	11.66 sec
실험 2	37/75	13.68 sec
실험 3	60/75	9.98 sec

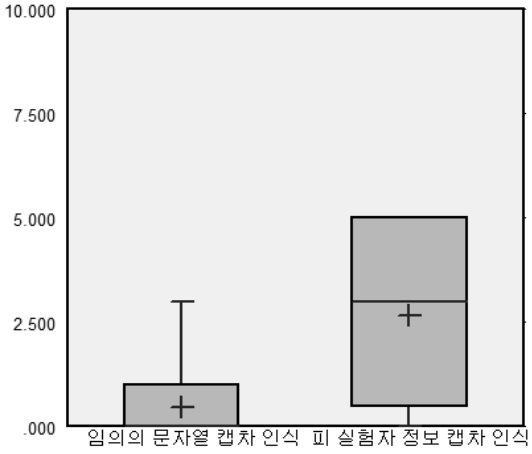
실험자가 피 실험자에게 제공하는 정보의 양은 인식률을 실험하기 위한 독립변인이다. 실험 1과 실험 3을 비교하면 인식률은 5.3%에서 80%로 캡차를 인식하는 평균 시간도 약 1.68초 빨라졌음을 알 수 있다. 실험 결과 피 실험자가 실험자로부터 캡차의 문자열을 힌트로 더 많이 제공받을수록 피 실험자의 캡차 인식률이 높아짐을 알 수 있었다. 분산 분석 결과 표본 평균간 유의 수준 $a=0.05$에서 유의한 차이가 있다는 결론을 얻었으며, 또한 Harada 등이 논문에서 밝힌 내용이 캡차에서 적용 가능함을 실험의 결과를 통해 알 수 있다.

〔표 2〕 왜곡의 정도의 따른 캡차 생성

				
단계 10	단계 9	단계 8	단계 7	단계 6
				
단계 5	단계 4	단계 3	단계 2	단계 1



(a) 피 실험자 정보 캡차 인식 실험



(b) 피 실험자 정보 캡차 인식 실험 결과

〔그림 4〕 실험자 정보 기반 캡차의 인식 정도

4.1.2 실험자 정보 기반 캡차 인식

다음은 사용자가 캡차의 문자열을 인지하고 있을 때와 그렇지 않은 경우에 캡차를 인식하는 정도를 알아보는 실험으로서, 실험 전 피 실험자로부터 몇 가지 정보를 입력받고 이 중 무작위로 하나를 선택한다. 그리고 선택된 피 실험자의 입력한 값과 동일한 길이의 임의의 문자열을 선택해 왜곡의 정도의 따라 〔표 2〕와 같이 10단계의 캡차를 만들고 왜곡을 줄여가면서 화면에 출력해주며, 피 실험자가 캡차를 인식하는 순간을 기록하였다.

실험의 관찰과 결과를 〔그림 4〕에 나타내었다. 〔그림 4-a〕 그래프의 세로 막대가 높을수록 어려운 캡차도 인식할 수 있다는 것을 보여준다. 세로막대가 없는 관찰은 단계 0을 가리킨다.

〔그림 4-b〕 결과로 보아 피 실험자가 정보를 인지하고 있었을 경우 그렇지 않았을 경우에 비해 더 어려운 캡차도 잘 인식할 수 있으며, 분산 분석 결과 표본 평균간 유의 수준 $a=0.05$에서 유의한 차이가 있다는 결론을 얻었다. 따라서 이를 인터넷 뱅킹에 적용한다는 가정 하에 사용자와 서버 간 송신자, 수신자, 계좌번호, 금액과 같은 사전 정보가 공유되고 이를 캡차 문자열로 사용한다면 캡차의 난이도가 높아도 사용자는 이를 유추해 낼 수 있음을 알 수 있다. 예를 들어 〔그림 5〕 수신자 이름인 "강전일"에 필터를 강도 높게 적용하여 '강x일' 로 보인다고 하더라도, 사용자는 수

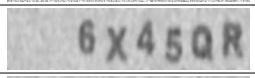
신자 이름이 '강전일' 인 사실을 미리 인지하고 있기 때문에 올바르게 선택할 수 있다는 것을 알 수 있다.

4.2 이용자 식별정도의 적절성

이용자 식별정도의 적절성을 위한 캡차의 난이도 설정은 두 가지 방법을 생각해 볼 수 있다. 첫째, 사용자가 캡차의 난이도를 직접 설정 하도록 하거나 둘째, 사용자에게 일방적으로 캡차를 제공하는 방법이다. 사용자가 직접 난이도를 설정하는 경우 계좌이체 전에 캡차 난이도 예시를 사용자에게 노출하고, 사용자로부터 난이도를 입력받아 사용자 개개인에게 해당하는 서비스를 제공함으로서 일방적으로 제공받는 방식보다 더 이용자 식별정도의 적절성을 만족할 수 있다.

캡차의 난이도는 캡차를 구성하는 방식에 따라 공격 성공률이 다르므로 어떤 방식의 캡차를 사용할 것인지, 그리고 캡차에 왜곡을 얼마나 주어야 하는지를 먼저 고려해야 한다.

〔표 3〕 가우시안 블러 왜곡 정도에 따른 캡차 생성

단계	왜곡의 정도	디캡차
원본 캡차		6X45QR
20		6X45QR
21		6X45QR
22		6X45QR
23		6X45QR
24		6X45QR
25		6X45QR
26		L741Q8
27		L711Q8
28		L711J8
29		L711J8
30		1711JF
35		1111JF
40		1111J1

이를 알아보기 위한 실험으로 97% 확률의 공격 성공률을 가지는 캡차와 디캡차 오픈소스 pwntcha을 사용하였다. 이미지 프로그램을 사용하여 캡차에 가우시안 블러 효과를 일정하게 주고 각각을 pwntcha를 이용해 디캡차하였다. 결과는 〔표 3〕에 20단계부터 정리하였다.

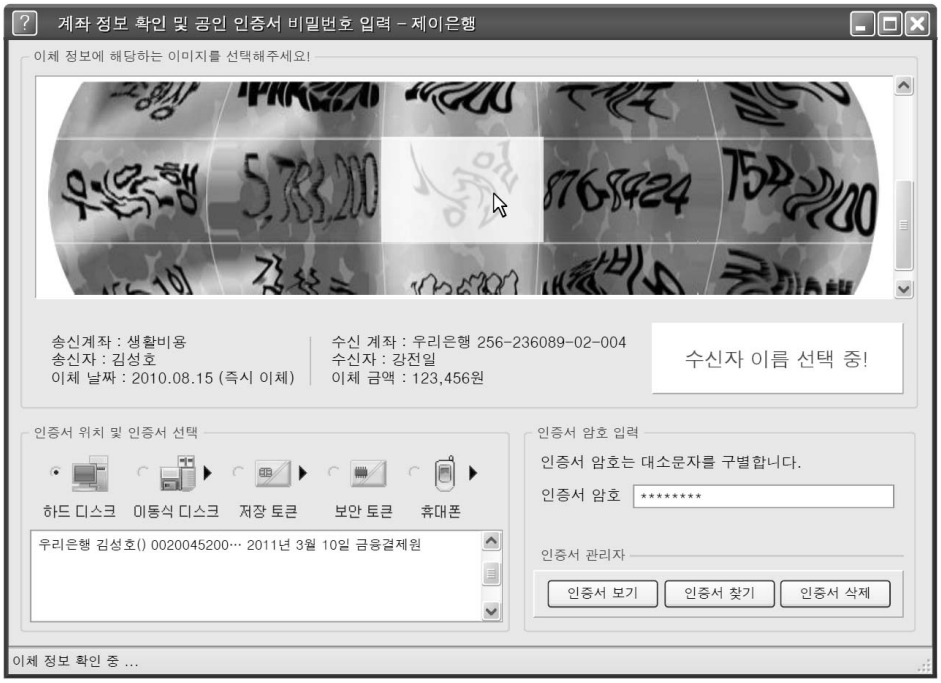
〔표 3〕에서 25단계 까지는 이미지에 왜곡을 주어도 캡차를 읽어내는 데에는 문제가 없었지만 26단계부터 왜곡에 의한 디캡차의 어려움이 있었다. 이 이후부터가 기존의 캡차를 보다 더 안전하게 사용할 수 있는 최소 왜곡(가우시안 블러 Radius: 2.6 pixels)의 정도이다.

실험에서는 보안에 취약한 캡차를 사용하였지만, 공격 성공률이 낮은 캡차를 사용해 왜곡을 줄수록 기존 캡차가 가지는 공격 성공률보다도 더 낮은 공격 성공률을 가지는 캡차를 얻을 수 있다.

캡차에 왜곡을 많이 줄수록 공격 성공률이 낮아지지만 과다한 왜곡이 적용될 경우 사용자조차도 식별할 수 없는 수준이 되어 튜링 테스트의 의미가 없어지는 점을 고려해야 할 것이다.

4.3 콘텐츠 기반 캡차를 이용한 승인 페이지 구성

〔그림 5〕는 콘텐츠(서버와 사용자간 알려진 정보) 기반 캡차를 이용하여 가상 이체 승인 페이지를 구성하였다. 이 화면은 계좌 이체 과정 중 보안카드 정보와 OTP 정보를 입력한 뒤 공인 인증서의 패스워드를 입력할 때 사용자에게 보일 것이다. 이 캡차는 사용자 정보 기반 캡차 6개(예를 들어 수신 계좌, 수신자, 수신 은행, 이체 금액, 송신자, 송신 계좌 등) 와 더미로 생성되는 44개의 캡차를 포함한 총 50개의 캡차로 구성되거나, 50개의 캡차를 병합하여 하나의 캡차로 구성할 수 있다. 단, 더미 캡차에 사용된 문자열이 사용자 정보 기반 캡차에 쓰인 문자열과 동일하다면, 사용자는 더미 캡차에서 올바른 캡차를 구별하기가 어려워질 것이다. 그러므로 더미 캡차에 쓰이는 문자열은 사용자 정보 기반 캡차에 사용된 문자열과 동일한 문자열을 선택하지 않도록 해야 하며 50개의 캡차는 문자열이 이미지로 렌더링 될 때 모든 문자열의 폭과 높이를 유사하게 렌더링 해야 할 것이다. 50개의 캡차가 한 화면에 보인다면 가독성이 저하될 수 있으므로 스



〔그림 5〕 콘텐츠 기반 캡차를 이용한 이체 승인 페이지의 예

크롤 하여 해당 캡차를 선택할 수 있게 구성하였다.

VPS와 MS 워터마크에서 캡차를 읽는다는 것은 서버에 유효한 OTP 문자열을 전송한다는 것을 의미하지만, 이 논문에서 제안하고 있는 콘텐츠 기반 캡차를 읽는다는 것은 서버에서 제시하는 올바른 계좌 이체 정보를 담고 있는 캡차를 선택할 수 있는지를 의미한다. 이 논문에서 제안한 기법을 사용할 때 악성 프로그램이 이체 정보를 변조할 경우, 변조된 계좌 이체 정보들이 캡차에 반영되기 때문에 사용자는 계좌 이체 정보가 변조되었음을 알 수 있다. 또한 우연히 사용자가 입력한 계좌 이체 정보가 더미 캡차로 생성된다면 사용자는 더미 캡차를 선택할 것이기 때문에 계좌 이체가 실행되지 않을 것이다. 이러한 사용자 정보 기반 캡차를 만들기 위해서 서버는 다음과 같은 작업을 수행해야 한다.

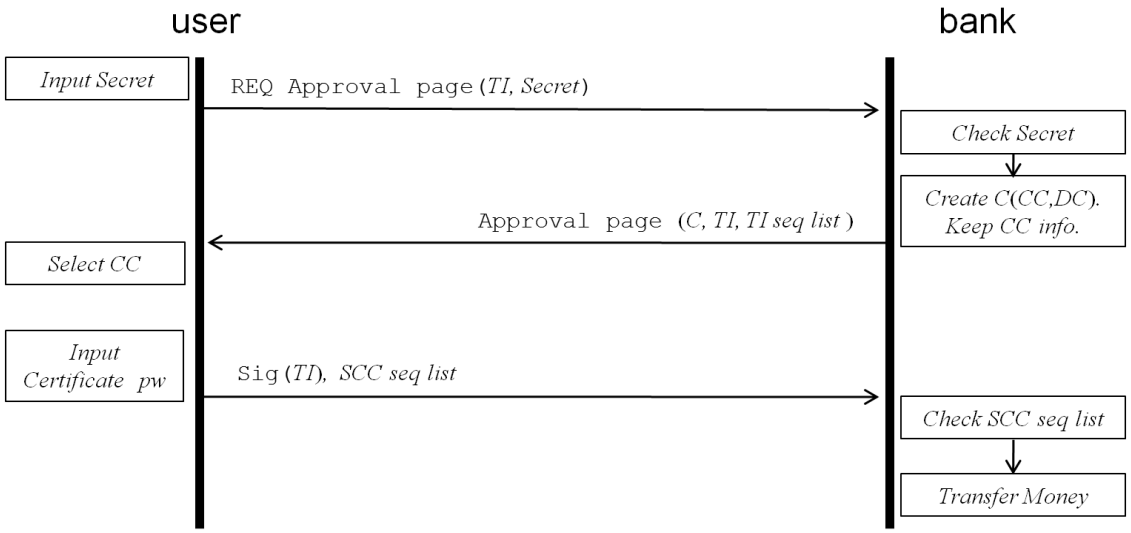
- 각각의 요소의 위치 정보 유지
  문자열이 렌더링 될 때, 서버에서는 캡차의 위치 정보를 저장해야 한다. 이 영역 정보들은 캡차가 새로 생성될 때마다 변경되며, 사용자가 서버로 전송한 좌표 값을 모두 검증할 때까지 유지해야한다.
- 계좌 이체 요소 제시 및 검증
  서버가 사용자에게 특정 계좌 이체 요소를 선택하는 문제를 제시하면, 사용자는 나열된 캡차에서 특정 좌표를 선택한다. 사용자가 선택한 좌표 정보를 전송받은 서버에서는 이 좌표 값이 문제로 제시한 이체 요소 영역에 해당하는 좌표인지 판단해야 한다. 캡차 요소들이 〔그림 5〕과 같이 직사각형으로 렌더링 되었다면, 서버는 영역의 시작 좌표(x, y)와 폭, 높이를 이용하여 사용자가 선택한 좌표가 서버가 제시한 해당 캡차의 영역 내에 포함되는지 여부를 검증할 수 있다.

4.4 콘텐츠 기반 캡차를 이용한 계좌 이체 승인과정

〔그림 6〕은 이 논문에서 제안한 기법을 사용할 때의 계좌 이체 승인 과정이다. 이전 과정들(서비스 로그인, 계좌 이체 요청 등)은 정상적으로 실행되었다고 가정한다.

4.4.1 승인 페이지 요청

사용자는 계좌 이체 정보 *TI*등을 기입한 후, 최종적으로 *Secret*값을 입력하여 서버에 승인 페이지를 요청한다. 이때 사용되는 *Secret*값은 보안카드 및 OTP 값으로 서비스 사용자를 인증하기 위해 사용된다.



〔그림 6〕 사용자 정보 기반 캡차를 이용한 계좌 이체 승인 과정

4.4.2 승인 페이지 전송

서버에서는 사용자로부터 전송된 *Secret*값에 대한 유효성을 검증한 후, *TI*를 이용하여 콘텐츠 기반 캡차 *CC* 및 더미 캡차 *DC*를 생성하고 이들을 병합하여 캡차 *C*를 만든다. 캡차 생성이 완료되면 서버는 사용자에게 캡차 *C*, 계좌 이체 정보 *TI*, 이체 정보 순서 리스트(*TI seq list*)가 포함된 승인 페이지를 전송한다. 한편, 서버에서는 계좌 이체 정보 순서 리스트(*TI seq list*)에 포함된 사용자 정보 기반 캡차 *CC*의 좌표 값(병합된 캡차 *C*에서 해당 사용자 정보 캡차 *CC*의 좌표 값)을 순서에 맞춰 저장하고 있어야 하며 이 좌표 값은 계좌 이체가 종료될 때 까지 서버는 유지해야 한다.

4.4.3 캡차 영역 선택 및 최종 계좌 이체 승인 요청

사용자는 〔그림 5〕처럼 구성된 승인 페이지에서 계좌 이체 정보 순서 리스트에 적혀있는 항목들을 순서대로 선택하여, 사용자 정보 기반 캡차 선택 리스트 *SCC*를 생성한다. 이 콘텐츠 기반 캡차 선택 리스트 *SCC*는 사용자가 선택한 항목들의 좌표 값을 담고 있다. 서버가 제시하는 캡차를 모두 선택한 후 사용자는 인증서 암호를 입력하고 계좌 이체 승인 버튼을 클릭하면, 공인인증서의 개인키로 계좌 이체 정보 *TI*를 서명하여 Sig(*TI*)를 생성하고 Sig(*TI*)와 콘텐츠 기반 캡차 선택 리스트 *SCC*를 서버로 전송한다. 이를 전달받은 서버에서는 콘텐츠 캡차 선택 리스트 *SCC* 값에 대한 유효성을 검증한 후 계좌 이체를 수행한다.

4.5 제안 기법의 안전성 분석

사용자 정보 기반 캡차 6개와 더미 캡차 44개로 구성된 승인 페이지에 다음과 같은 공격에 대한 안전성을 분석한다.

4.5.1 유사 VPS에 대한 모의 공격

이 논문에서 제시한 기법이 기존의 방법과 비교해서 얼마다 더 안전성이 있는지 비교하기 위해 Arcot의 VPS와 유사한 캡차를 제작하고 OTP를 추출하는 실험을 진행하였다.

〔표 4〕 유사 VPS 캡차에 대한 OTP 추출 실험 결과

	생성 캡차 수	추출 성공	추출 실패	성공 확률
실험 1	1000	823	177	82.3 %
실험 2	1001	837	164	83.6 %
실험 3	1005	823	182	81.9 %

실험에 대한 결과를 〔표 4〕에 기록하였다. 세 번의 실험을 통한 결과 모두 80% 이상의 확률로 VPS의 OTP를 추출할 수 있었다.

4.5.2 무차별 선택을 통한 공격

악성 프로그램이 문자열을 추출해내지 못하지만, 유사한 캡차를 제작할 수 있다고 가정 하에 악성 프로그램은 50개의 캡차중 무작위로 6개의 캡차를 선택하여 승인 페이지를 변조하는 공격을 실행 할 수 있다. 이 때의 공격 성공률은 $1/(50\times\cdots\times45)=8.74\times10^{-11}$이다.

4.5.3 캡차를 변조하는 공격

특정 문자열을 추출하여 악성 프로그램이 캡차를 변조하는 공격은 다음과 같이 두 가지로 나누어진다.

- 캡차를 요소별로 분리하는 과정에 대한 안전성 분석
  콘텐츠 기반 캡차의 경우, 필터의 적용 강도를 기존의 방법들보다 높일 수 있고 요소별 분리를 방해하는 필터들을 적극적으로 사용할 수 있는 장점이 캡차의 요소들을 서로 분리하는 공격에 적극적으로 대응할 수 있다. MS 워터마크는 OTP 문자열 값을 반복적으로 사용하고 위치 또한 고정되어 있다. OTP 문자열에 필터 적용 강도를 높이더라도 악성 프로그램은 여전히 OTP문자열을 분리해낼 수 있으며, 사용자는 OTP 문자열을 인식하기 어렵기 때문에 서버가 보낸 OTP 문자열이 자신의 OTP 생성 기기로부터 생성된 OTP 문자열과 일치하는지 여부를 판단하기 쉽지 않다는 문제점이 여전히 존재한다. 이 논문에서 제안한 기법은 앞선 문제점들을 개선하였기 때문에 MS 워터마크보다 높은 안정성을 가진다.
- 변조한 정보를 찾아 사용자가 입력한 정보로 다시 변조하는 공격에 대한 안전성 분석
  VPS와 MS 워터마크는 OTP 문자열만 분리해낼 수 있다면 악성 프로그램의 의도대로 캡차를 변조 할 수 있다. 하지만 이 논문에서 고안한 기법에서는 올바른 계좌 이체 정보 요소들 중에서 악성 프로그램에 의해 변조된 정보 요소들을 찾아 이를 사용자가 입력한 정보들로 변조해야 한다. 제안한 기법에서 출력하는 50개의 캡차 중에서 최대 6개(송·수신 계좌, 송신자, 수신자, 수신 은행, 이체 금액)의 사용자 정보 기반 캡차를 찾아 바꿔야 한다. VPS의 경우 화면에 4개의 OTP를 포함한 4개의 계좌 이체 정보만 출력하기 때문에 0.25의 확률로 공격에 성공할 것이다. 하지만 이 논문에서 제안한 기법의 경우에는 이 같은 공격이 성공할 확률은 $1/(50\times49)=4.08\times10^{-4}\sim1/(50\times\cdots\times46)=3.93\times10^{-9}$로 VPS에 비해 상당히 낮은 공격 확률을 보인다.

V. 결론 및 향후 연구

이 논문에서는 사용자와 서버 사이에 알려진 계좌 이체 정보를 이용하여 콘텐츠 기반 캡차를 제안하였다. 또한 계좌 이체 정보에 해당되지 않는 요소들을 더미 캡차로 추가하여 악성 프로그램이 선택해야 할 대상을 늘림으로서 공격에 성공할 확률을 낮추었다. 악성 프로그램이 인공지능 및 이미지 프로세싱 기술 등을 이용해 캡차 문자열을 읽어 승인 페이지를 변조하는 공격에 대해서도 다양한 필터를 강도 높게 적용하는 방식을 통해 기존 방법들보다도 유연하게 대응할 수 있다는 장점을 가진다.

더불어 자동화 프로그램에 의해 실행되는 공격에 적극적으로 대응할 수 있기 때문에 수많은 단말기가 동시에 공격당하는 것을 미연에 방지하여 피해규모를 줄일 수 있을 것이다. 향후 연구과제로는 실험을 통해 다수의 캡차 생성이 서버의 성능에 얼마나 영향을 미치는지를 살펴보고 캡차의 개수를 적절하게 조정해야 한다. 더불어 모바일 뱅킹이 가능한 단말기에도 적용이 가능한지 여부를 살펴 봐야할 것이며, 사용자에 따른 식별정도의 적절성을 올바르게 제시할 수 있어야 할 것이다.

참고문헌

[1] L.V. Ahn, M. Blum, N.J. Hopper, and J. Langford, "CAPTCHA: telling humans and computers apart," Euro-crypt 03, LNCS 2656, pp. 294-311, May. 2003

[2] R.A. Gopalakrishna, "Authentication using a turing test to block automated attacks," US 2009/0199272 A1, US Patent, Aug. 2009

[3] D.J. Steeves and M.W. Snyder, "Secure online transactions using a CAPTCHA image as a watermark," US 2007/0005500 A1, US Patent, Jan. 2007.

[4] Finjan Malicious Code Research Center,

"Cybercriminals use trojans & money mules to rob online banking accounts," http://www.finjan.com/getobject.aspx?objid=679

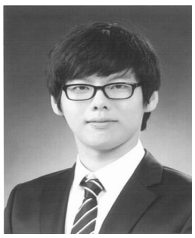
〔5〕 P. Guhring, "Concepts against man-in-the-browser attacks," http://www2.futureware.at/svn/sourcerer/CAcert/SecureClient.pdf

〔6〕 맹영재, 신동오, 김성호, 양대헌, 이문규 "국내 인터넷뱅킹 계좌이체에 대한 MITB 취약점 분석," Internet and Information Security, 1(2), pp. 101-118, 2010년 11월.

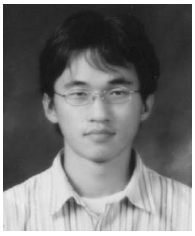
〔7〕 S.Y. Huang, Y.K. Lee, G. Bell, and Z.h Ou, "A projection based segmentation algorithm for breaking MSN and YAHOO CAPTCHAs," In Proc. of the 2008 International Conference of Signal and Image Engineering, pp. 727-730, July. 2008.

〔8〕 A. Harada, T. Isarida, T. Mizuno, and M. Nishigaki, "A user authentication system using schema of visual memory," In Proc. of Bio-ADIT'06, pp. 338-345. Jan. 2006.

〈저자소개〉



이 상 호 (Sang-ho Lee) 학생회원
2011년 2월: 공주대학교 정보통신공학과 졸업
2013년 3월~현재: 인하대학교 컴퓨터정보공학과 석사과정
〈관심분야〉 시스템 보안, 웹 보안



김 성 호 (Sung-ho Lee) 학생회원
2009년 2월: 인하대학교 컴퓨터 공학과 졸업
2011년 7월: 인하대학교 정보통신공학과 석사
2011년 8월~현재: 국가보안기술연구소 연구원
〈관심분야〉 시스템 보안, 네트워크 보안

강 전 일 (Jeon-il Kang) 학생회원
2003년 2월: 인하대학교 컴퓨터 공학과 졸업
2006년 2월: 인하대학교 정보통신대학원 석사
2006년 3월~현재: 인하대학교 정보공학과 박사과정
〈관심분야〉 RFID 보안, 생체 인식 보안, WSN 보안, 무선 인터넷 보안, 웹 인증 보안

변 제 성 (Je-sung Byun) 학생회원
2009년 2월: 공주대학교 컴퓨터 공학과 졸업
2010년 3월~현재: 인하대학교 컴퓨터정보공학과 석사과정
〈관심분야〉 무선 센서 네트워크 보안, 무선 인터넷 보안, 인증 프로토콜, 웹 인증 보안

양 대 헌 (Dae-hun Nyang) 정회원
1994년 2월: 한국과학기술원 과학기술 대학 전기 및 전자 공학과 졸업
1996년 2월: 연세대학교 컴퓨터 과학과 석사
2000년 8월: 연세대학교 컴퓨터 과학과 박사
2000년 9월~2003년 2월: 한국전자통신연구원
정보보호연구본부 선임연구원 2003년 2월~현재: 인하대학교 컴퓨터정보공학과 부교수
〈관심분야〉 암호이론, 암호프로토콜, 인증프로토콜, 무선 인터넷 보안

이 경 희 (Kyung-hee Lee) 정회원
1998년 8월: 연세대학교 컴퓨터 과학과 석사
2004년 2월: 연세대학교 컴퓨터 과학과 박사
1993년 1월~1996년 5월: LG소프트(주) 연구원
2000년 12월~2005년 2월: 한국전자통신연구원 선임연구원
2005년 3월~현재: 수원대학교 전기공학과 조교수
〈관심분야〉 바이오인식, 정보보호, 컴퓨터비전, 인공지능, 패턴인식