보안공학연구논문지 *Journal of Security Engineering*
Vol.10, No.6 (2013), pp.655-668
http://dx.doi.org/10.14257/jse.2013.12.01

SVM을 이용한 스미싱 탐지기법

이지원[1], 이동훈[2], 김인석[3]

Method of Detecting SmiShing using SVM

Ji-Won Lee[1], Dong-Hoon Lee[2], In-Suk Kim[3]

요 약

IT기술의 발전으로 다양한 단말기를 통해 전자금융거래가 이루어지고 있다. 특히 스마트폰 사용이 급증함에 따라 무선기기를 이용한 금융거래가 대중화되면서, 이를 이용한 금융사기가 심각한 사회문제로 대두되고 있다. 이런 사기 피해에 대해 여러 기관들(수사기관, 금융기관, 통신사 등)에서 많은 보안대책을 제시하고 있으나, 피싱 수법의 계속적인 진화로 인해 피해방지에 어려움을 겪고 있다. 이렇게 진화하는 피싱 기법들 중 특히 사회 공학적 기법을 이용한 스미싱으로 인한 피해 증가가 두드러지고 있는 추세이다.

본 논문에서는 스미싱에 대한 기존 탐지방법들과 한계를 분석한다. 이를 바탕으로 문자 메시지 내용과 발신자 주소 특성에 초점을 맞춘 기계학습 알고리즘을 사용하여 스미싱 문자 메시지와 일반 문자 메시지를 분류하는 방법을 제안한다. 또한 실험을 통해 제안하는 방법이 기존 연구들보다 스미싱 문자 메시지 분류에 더 높은 정확도를 제공함을 보인다.

핵심어 : 스미싱, SVM, 기계학습, 스마트폰

Abstract

As IT technologies develop, electronic financial transactions are performed via the various devices. Financial transactions using a wireless device especially by use of a smart-phone is rapidly increasing, financial fraud using the smart-phone has become a serious social issue. Although various organizations such as investigative authority, financial institution, etc. present a lot of security measures, they have a hard time to prevent by the continuing evolution of phishing. In particular, smishing using social engineering skills is outstanding phishing technique recently.

In this paper, we analyze existing detection methods and its limitations for smishing. Based on these works, we propose classification method using a machine learning algorithm focused on characteristics of message contents and sender's address. In addition, we demonstrate that accuracy of the proposed method are better than other methods with the experimental results.

Keywords : smishing, SVM, machine learning, smartphone

접수일(2013년10월15일), 심사의뢰일(2013년10월16일), 심사완료일(1차:2013년10월30일, 2차:2013년11월11일)
게재일(2013년12월31일)

[1]136-701 서울시 성북구 안암동 111, 고려대학교 금융보안학과 석사과정.
email: miruesj@naver.com
[2]136-701 서울시 성북구 안암동 111, 고려대학교 정보보호대학원 교수.
email: donghlee@korea.ac.kr
[3](교신저자) 136-701 서울시 성북구 안암동 111, 고려대학교 정보보호대학원 교수.
email: iskim11@korea.ac.kr

* 본 연구는 미래창조과학부 및 한국인터넷진흥원의 "고용계약형 지식정보보안 석사과정 지원사업"의 연구결과로 수행되었음(과제번호 H2101-13-1001)

ISSN: 1738-7531 JSE

1. 서론

2006년 6월 보이스 피싱이 처음 등장한 이후, 피싱에 대한 피해현황이 매년 증가하고 있다. 2013년 6월말 기준으로 전체 스마트 폰 기반 모바일뱅킹 등록 고객 수는 3,131만 명으로 2012년 12월말 2,087만 명보다 11.5% 증가하였으며, 특히 스마트 폰을 이용한 모바일뱅킹 서비스 이용건수는 2,032만 건, 금액은 1조 3,523억 원으로 전체 모바일뱅킹 이용건수 및 금액의 97%이상을 차지하고 있다. 이에 따라 스미싱을 이용한 모바일 뱅킹 피해도 크게 우려되고 있는 상황이다[12].

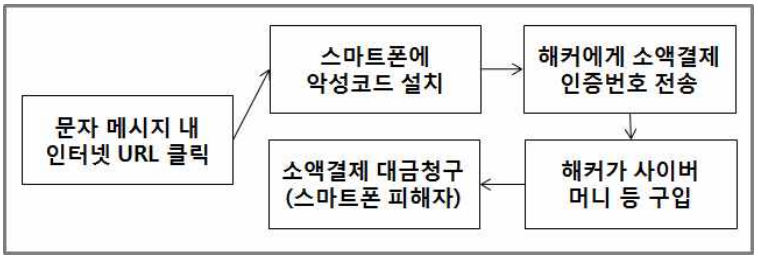
스미싱(SmiShing)이란 SMS(문자메시지)와 Phishing의 합성어로, 문자메시지를 이용한 피싱을 의미한다. 주로 단축 URL을 이용하여 악성 앱 설치를 유도한 후, 개인비밀정보를 요구하거나 스마트 폰 소액결제 서비스를 악용하는 신종사기 수법이다. 스미싱의 내용은 [그림1]과 같이 금융기관 사칭, 정부지원금 환급 사칭, 무료쿠폰 지급, 이메일 명세서, 모바일 청첩장 등 다양한 내용들로 수신되고 있다.

[그림 1] 스미싱 문자 메시지

[Fig. 1] Smishing text message

스미싱 사기에 대한 절차를 보면 [그림2]와 같다. URL이 포함된 문자 메시지가 수신되었을 경우, 해당 URL을 클릭하면 해커가 만든 서버에 접속하여 스마트 폰에 악성코드를 다운로드한 후 설치를 유도한다. 그 후 해커는 해당 앱을 통해 이용자의 개인정보나 금융정보를 탈취하여 소액결제 대금 청구 등을 하게 된다. 스미싱의 피해사례를 통해 스미싱의 행위를 분석해 보면 메시지 내용은 다양하지만 공통된 부분은 URL이 포함되어 있는 문자 메시지를 선택했을 경우 악성코드가 설치된다는 점이 공통적으로 나타난다. 여러 기관에서 악성 URL을 Blacklist로 관리하는 대응책을 내놓고 있지만, 기존 대응책은 스미싱 문자 메시지를 예방하는데 한계가 있다. 스미싱 문자 메시지에 포함된 URL이 자주 변경되고 단축 URL로 구성되어 있어서 유해한 사이트인지 구별하기 힘들

고 Blacklist를 우회할 수도 있기 때문이다[4].



[그림 2] 스미싱 사기에 대한 절차
[Fig. 2] Procedure of Smishing Fraud

본 논문에서는 형태소 분리기를 통해 문자 메시지로부터 명사를 추출하여 스미싱에서 사용되는 단어들이 있는지를 확인하고, 본 논문에서 제안하는 스미싱 메시지의 특성들을 이용하여 SVM을 이용한 유사도 검사를 통해 수신된 문자 메시지가 스미싱 문자인지 일반 문자인지를 판별하여 사용자에게 고지할 수 있는 방안을 제안한다. 본 논문의 구성은 다음과 같다. 2장에서는 명사를 추출하는 방법과 SVM에 대하여 알아보고 3장 및 4장에서는 스미싱 탐지를 위한 시스템 구조 및 방법은 제안하다. 4장에서는 실험을 통해 본 논문에서 제안하는 기법의 효율성을 설명하고, 5장에서는 결론을 맺는다.

2. 배경지식

2.1 명사추출기

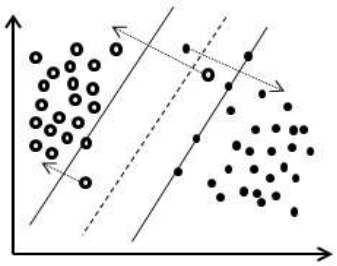
명사를 추출하는 방법에는 형태소 분석기를 이용하는 방법, 형태소 분석기와 품사 태거를 이용하는 방법, 언어분석 도구를 사용하지 않는 방법 등 크게 세 가지로 나눌 수 있다[13].

형태소 분석기를 이용하는 방법은 형태소 분석기를 통해 명사를 추출하는 방식으로 가장 보편적으로 사용되는데 명사 추출에 대한 정확도는 높은 편이지만, 중의성 문제에 따라 성능의 차이가 나타날 수 있다. 형태소 분석기와 품사 태거를 이용하는 방법은 형태소 분석기를 이용하는 방법과 마찬가지로 명사 추출에 대한 정확도가 높고 중의성 문제도 품사 태거를 통해 해결할 수 있으나, 분석 속도가 느리고 문장을 분리하는 전처리 과정이 필요하다. 언어분석 도구를 사용하지 않는 방법은 형태소 분석기와 같은 언어 분석 도구를 사용하지 않아 사전의 의존도가 크고 미등록어 처리가 어렵지만, 분석이 빠르며 모바일 장치와 같이 제한적인 자원을 사용하는 환경에서 사용이 용이하다[14]. 본 논문에서 제안하는 방법은 스마트 폰을 기반으로 하기 때문에 형태소 분석기와 같은 언어 분석 도구를 사용하지 않는 방법을 선택할 수도 있으나, 문자 메시지는 책과 다르게 문법이

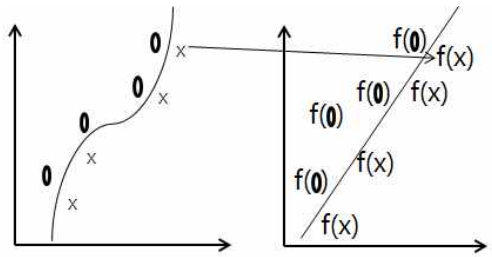
나 어휘가 틀린 문구를 많이 포함하고 있어 미등록어 처리가 어려운 면이 있다. 따라서 기분석 사전을 추가하여 문장 단위로 규칙을 찾아 미등록어의 오분석을 줄이고자 하였다.

2.2 SVM원리

SVM(Support Vector machine)은 1992년 Boser 등에 의해 제안된 기계학습 알고리즘이다. 이는 학습 데이터를 통해 데이터를 이분법으로 분리하는 초평면 중에서, 데이터들과 가장 거리가 먼 초평면을 찾는 방법이다. 커널 매핑 개념과 최적화 기술을 통계적 학습의 원리에 통합한 알고리즘으로 초평면을 찾은 후에는 입력된 벡터 값의 좌표의 위치에 따라 해당 클래스를 반환하게 된다. SVM은 이분법으로 선형 분리가 가능하지만 실제 데이터를 분리하는 경우에는 선형 분리가 불가능한 경우가 많이 존재한다. 이러한 경우 다음 [그림 3]과 같이 슬랙변수를 적용하거나 혹은 [그림4] 와 같이 커널함수를 적용하여 이분법으로 만들어주게 된다 [8][9]. 커널함수로는 RBF 커널, Poly 커널 등이 있으며, 입력 공간의 비선형적인 높은 차수를 선형적으로 해석할 수 있다.



[그림3] 슬랙변수를 적용한 SVM
[Fig. 3] SVM Slack variable



[그림4] 커널을 사용하여 이분법으로 만든 SVM
[Fig. 4] SVM using Kernel Function

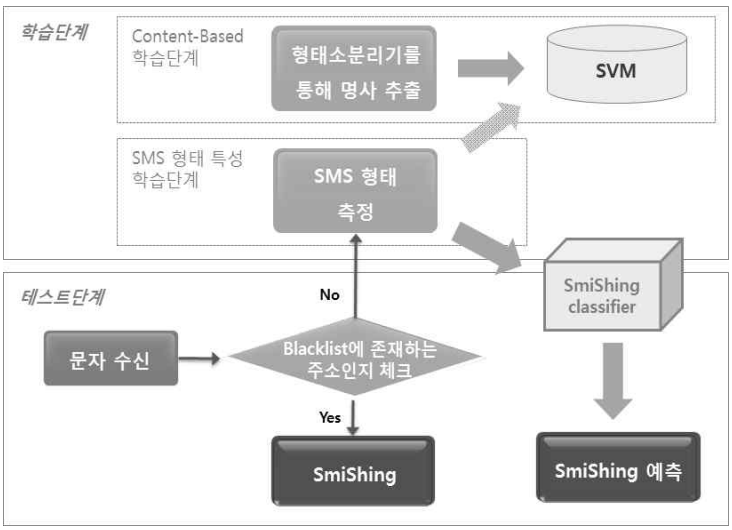
3. 제안하는 방법

본 논문에서는 스미싱과 일반문자를 구분하기 위해 SVM을 통한 분류기 방법을 제안한다. 스미싱 문자 메시지나 일반 문자 메시지의 특성들은 유사한 점이 많기 때문에 스미싱 문자의 특성을 어떻게 선택하느냐에 따라 분류 율이 크게 달라질 수 있다.

본 논문에서 제안하는 스미싱 분류를 위한 시스템은 학습단계와 테스트 단계로 나뉜다. 학습단계는 Content-Based 학습단계와 SMS 형태 특성 학습단계로 나눈다. [그림5]은 본 논문의 스미싱 분류를 위한 시스템 구조를 나타낸다. 시스템이 동작하는 방식은 다음과 같다.

Content-Based 학습단계에서는 수신된 문자에서 형태소 분리기를 통해 추출한 명사가 스미싱 메시지에서 자주 나오는 단어인지를 학습을 하고 빈도수를 데이터로 가지고 있다. 이 데이터는 SMS 형태 특성중 하나이다. SMS 형태 특성 학습 단계에서는 SMS 형태 특성을 측정하여 스미싱

메시지인지 일반 메시지인지를 분류하기 위해 SVM 알고리즘을 통해 학습을 시킨다. 학습단계를 거친 후 테스트 단계를 수행하게 된다. 테스트 단계에서는 문자가 수신되면 Blacklist 데이터베이스에 존재하는 전화번호인지를 검사한다. Blacklist 데이터베이스에 존재하는 전화번호일 경우 스미싱 메시지임을 사용자에게 알려주며 존재하지 않는 전화번호일 경우 SVM 알고리즘을 이용한 학습을 통해 스미싱 분류를 수행하여 스미싱 메시지임을 예측하도록 한다.



[그림 5] 스미싱 분류를 위한 구조
[Fig. 5] Structure of Smishing Classifier

3.1 Blacklist 기법의 주소 검사방법

주소를 검사하는 방법에는 Blacklist 기반 기법과 Whitelist 기반 기법이 있다. Whitelist 기반기법은 수신을 받은 번호가 Whitelist에 포함되어 있을 경우 안전한 번호라 인지하고 사용자에게 일반문자 메시지로 알려주게 된다. Whitelist에는 신뢰할 수 있는 주소만 저장되기 때문에 목록에 존재하지 않을 경우 신뢰도가 높은 번호라 하더라도 문자 메시지를 수신하지 못하게 되는 단점이 있다. Blacklist기법은 사용자가 수신을 받은 번호가 Blacklist에 포함되어 있을 경우 사용자에게 피싱과 관련된 번호라는 것을 알려주게 된다. Blacklist 기법은 오탐률이 적고 구현이 쉬워 보편적으로 사용된다. 본 논문에서도 Blacklist기법을 이용하여 우선적으로 스미싱 문자 메시지를 판별한다[5].

Blacklist에 등록된 데이터의 기준은 휴대폰 사용자들이 수신 받은 문자 메시지에서 피해를 보거나 스미싱 문자 메시지로 의심하는 문자 메시지에 대해 포털 사이트에 등록한 것을 기준으로 했다. 포털 사이트에서는 스미싱 문자 메시지뿐만 아니라 피싱 데이터들에 대한 정보를 가지고 있으며, 하루에도 수십 건씩 스팸 번호가 등록되고 있으므로, 포털 사이트에 있는 데이터들을 주기적으

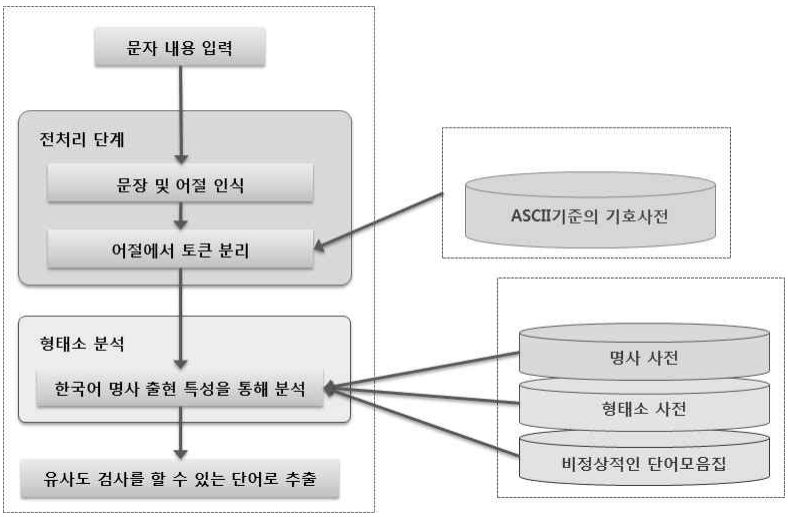
로 업데이트함으로써, Blacklist 기법으로 스미싱 문자 메시지를 되도록 많이 감지하도록 하였다.

3.2 형태소 분리를 이용한 명사 추출

Blacklist 기법을 이용하여 추출되지 않은 문자 메시지들은 기계학습과 유사도 측정을 하여 스미싱 여부를 확인하기 위해 수신된 문자메시지로부터 명사를 추출한다. 명사를 추출하는 방법 중 형태소 분석기를 이용하는 방법, 형태소 분석기와 품사 태거를 이용하는 방법 등이 있으나 본 논문에서는 기분석 사전을 이용하여 형태소를 분리하여 명사를 추출한다.

새로운 문자 메시지가 수신되면 전처리 단계를 거치게 된다. 일반적으로 스미싱 문자 메시지의 경우 "인출", "대출", "공짜"와 같은 단어가 content based filtering에서 검출되는 것을 막기 위해 글자 중간에 아스키 코드를 삽입하여 검출되기 어렵게 되어있다. 우선 ASCII기준의 기호사전을 이용하여 어절이 형태소별로 구분하기 좋은 구조로 재구성하고, 형태소 분석기를 이용하여 명사를 추출한다.

수신된 문자 메시지를 확인해 보면 일반문자 메시지나 스미싱은 한국어 어절에 맞지 않고 축약형 문자 메시지나 이상한 형태로 단어를 형성함으로써, 형태소 분리기를 통해서 단어를 추출할 경우 명사 추출 정확도가 일반 문장보다 떨어진다. 따라서 스미싱의 탐지율을 높이기 위해서 정확한 스미싱 단어를 추출하는 것이 중요하다. 본 논문에서는 [그림 6]와 같이 형태소분리기를 구성했다.



[그림 6] 형태소 분리기 구성

[Fig. 6] Structure of Morpheme Analyzer

$$\text{정확률}(\%) = \frac{\text{올바르게 인식한 명사의 개수}}{\text{인식된 전체 명사의 개수}} \times 100 \quad (1)$$

정확한 단어를 추출하기 위해 비정상적인 단어 모음집을 구성하여 스미싱 문자 메시지에서 자주 사용하는 패턴을 탐지하도록 구성하였다. 비정상적인 단어 모음집을 추가함으로서 스미싱 단어를 추출하는 확률이 높아졌음을 보여주고 있다. 정확률은 올바르게 인식한 명사의 개수를 형태소에서 분류하여 인식된 전체 명사의 개수로 나눔으로서 정확률을 계산하였다. 200개의 스미싱 문자 메시지를 형태소를 통해 명사를 추출한 결과이다. 스미싱 문자 메시지의 경우 명사중간에 ASCII코드를 삽입하여 명사를 걸러내기 힘들다. 예를 들어 "농협"이란 단어가 스미싱일 경우 "농/협" 이나 "농.협."으로 수신된다. 그렇기 때문에 명사를 추출해 내기가 일반 문장에 비해 힘들다. 전처리 단계에서 ASCII값이 0x80 이하의 값 중에서 기본적인 기호값 들은 제거를 하고 형태소를 분리할 수 있도록 어절을 최적화 하였다. 더 정확한 스미싱 단어를 추출하기 위해 비정상적인 단어 모음집인 데이터 베이스를 생성하여 형태소 분리기와 전처리 단계에서는 추출되지 못하는 단어들을 저장하도록 하였다. 분석을 통해 정확률은 (1)과 같이 계산된다.

3.3 스미싱 탐지 기준

스미싱을 탐지하기 위해 문자 메시지를 구성하는 SMS형태와 문자 메시지 안에 URL이 있을 경우 URL의 특성, 방송통신위원회에서 규정된 전화번호 체계, Web site 정보, 형태소 분리기, 기타의 특성을 가지고 효율적으로 스미싱 문자 메시지를 탐지한다. [표1]는 스미싱 문자 메시지와 일반문자 메시지의 특징을 분석하여 속성을 탐지 기준으로 구성한 것이다[1][2][3].

3.3.1 SMS 형태

SMS 형태는 수신을 받은 메시지의 가장 기본적인 속성을 가지고 스미싱을 판별한다. 수집한 자료를 분석한 결과 스미싱일 경우 MMS인 경우보다 SMS일 확률이 현저히 높았으며 문자 메시지 내용 중에 URL이나 번호들이 포함하고 있으며, 문자 메시지 내용 안에 포함되는 이모티콘의 개수도 현저히 높았다. 이모티콘의 경우 ASCII코드의 값이 0x80이상인 확장 이모티콘을 일컫는다.

[표 1] 스미싱 탐지 기준
[Table 1] Detection criteria of Smishing

대분류	소분류	내용
SMS 형태	메시지 형태	SMS, MMS 여부
	메시지 내용	문자 메시지 내용 중에 URL이나 번호가 있는지
	이모티콘 수	문자 메시지내용 중에 이모티콘 개수
발신 전화번호 형태	전화번호 체계	방송통신위원회 고시 전기통신번호관리세칙에 의해 규정된 대한민국 전화번호 체계에 대한 수신자 번호 길이 오류 여부
Web site 정보에 대한 형태	apk 존재	HTML내의 Header tag 분석을 통해 apk 다운로드 받는지 분석

문자내용 빈도수	명사 빈도수	스미싱에 자주 사용되는 단어의 빈도수 측정
기타	도메인 비교	대표적인 회사에 대한 등록되어 있는 전화번호와 회사 도메인 이름 대조
	수신시간 비교	문자 메시지 수신 시간과 내용을 비교

3.3.2 발신 전화번호 형태

전화번호 형태는 방송통신위원회 고시 전기통신관리세칙에 의해 규정된 대한민국 전화번호 체계가 구축되어 있다. 예를 들어 전화번호 체계를 보면 휴대폰일 경우 01X-YYYY-ZZZZ 형식으로 되어 있으며 YYYY는 최대 4자리이고 010일 경우 4자리로 구성되어 있으며 맨 첫 자리는 2부터 9까지의 숫자로 시작한다. 스미싱 문자 메시지를 보면 010으로 시작하지만 YYYY가 3자리일 경우도 있으며, YYYY의 첫째자리가 1로 시작하는 경우는 활성화된 전화번호가 아니다.

3.3.3 Web site 정보에 대한 형태

수신된 문자 메시지 중에 URL이 포함되어 있을 경우 사용자들이 URL을 클릭할 경우 apk가 다운로드 되어 사용자가 모르는 사이에 스마트 폰에 설치되는 경우가 있다. 본 논문에서는 URL을 접속하여 HTML내의 Header tag 분석을 통해 apk를 다운로드 받는지를 분석하였다. 다른 스미싱 탐지 앱들은 URL에 들어가서 apk를 다운받는지를 체크하지만, 본 논문에서는 Header의 정보를 분석하여 Content-disposition을 포함하고 있는지를 확인하여 apk를 다운로드 받는지를 체크한다. Content-disposition이 포함하고 있을 경우는 다운로드할 파일을 가지고 있는 경우로서 네트워크상으로 body를 다운받기 전에 미리 차단을 시킬 수 있다.

3.3.4 문자내용 빈도수

형태소 분리기를 통해 추출한 명사들의 빈도수를 검사하여 스미싱 문자 메시지에서 자주 나타나는 단어들을 SVM을 통해 벡터화 작업을 통해 스미싱 데이터베이스로 가지고 있다. 문자 메시지가 수신되면 형태소 분리기를 통해 명사를 추출하여 스미싱 데이터 베이스와 비교하여 스미싱 메시지와 관련된 단어의 개수를 확인한다.

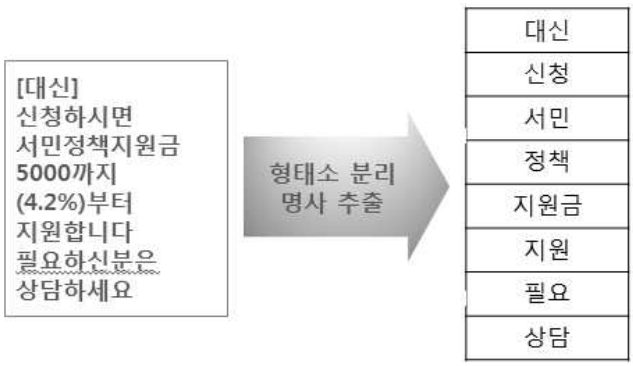
3.3.5 기타

기존에 한국에서 대표하는 기관들의 전화번호와 도메인 데이터베이스를 가지고 있으며, URL이나 전화번호를 데이터베이스와 비교하도록 한다. 스미싱 문자 메시지의 경우 문자 메시지 수신시간이 오후와 새벽 시간대에 많이 수신되며 내용 또한 스미싱과 연관된 단어들이 포함된 문자 메시지가 수신되는 것을 볼 수 있다.

4. 실험환경 및 실험결과 분석

4.1 실험환경

본 실험에서는 안드로이드 운영체제로 4.1.2버전, Galaxy Note II와 Java application을 기반으로 Eclipse Editor를 이용해 개발을 하고 테스트를 하였다. 데이터를 수집한 방법은 2013년 5월부터 2013년 6월까지 포탈 사이트상에서 찾은 스미싱 문자 메시지들, 지인들을 통해 받은 스미싱 문자 메시지와 일반문자 메시지(광고문자 메시지, 전화번호부에 저장되어 있는 사람의 문자 메시지)를 수집하였다. 중복 메시지를 제외하고 300개의 문자를 수집하여, 총600개로 실험 데이터를 사용하였다.



[그림 7] 형태소 분리기 결과값
[Fig. 7] Result of Morpheme Analyzer

형태소 분리기는 공개되어 있는 Java용 라이브러리를 기반으로 개발하였으며 형태소를 분리하기에 앞서 스미싱 문자 메시지는 다양한 형태로 들어오기 때문에 바로 형태소 분석기를 돌렸을 경우 정확도가 떨어지게 되므로 전처리 과정을 거치도록 하였다. ASCII 사전을 통해 전처리 과정을 거쳐서 형태소 분석기에 돌리기 적합한 형태로 구성 후 형태소 분리기를 사용하였다. 전처기 과정과 비정상적인 단어 모음집을 추가하여 형태소 분리기를 개발을 하였으며 수신된 문자 메시지를 형태소 분리시킨 결과 값의 형태는 [그림7]과 같다.

[표2]은 세 가지로 구분하여 정확률을 계산한 수치이다. '공ㅠr', '출금', 등과 같은 단어들을 비정상적인 단어 모음집에 저장함으로써 명사를 추출하는 정확률은 [형태소], [형태소 + 전처리 단계]를 거친 것 보다 훨씬 높아진 것을 볼 수 있다.

[표 2] 형태소 분리기의 정확률 결과

[Table 2] Precision results of Morpheme Analyzer

	정확률
형태소	72.39
형태소 + 전처리 단계	91.05
형태소 + 전처리 단계 + 비정상적인 단어모음집	94.97

형태소 분리기를 통해 추출된 명사들을 가지고 빈도수를 체크하도록 한다. 빈도수가 높은 명사들은 스미싱 문자 메시지에서 사용이 빈번한 명사로 저장하여 단어의 빈도수를 검사하여 스미싱 메시지과 일반메시지를 구별하도록 한다.

```
case 1 "Mssage Type"
        SMS : 1, MMS : 2
case 2 "Message Attribute"
        TRUE, FALSE  /* 문자 메시지 내용중에 URL이나 번호가 있을 경우 TRUE */
case 3 "Number of Emoticons"
        INT  /* 0x80 이상값을 가지고 있는 확장 ASCII에 있는 이모티콘 */
case 4 "Phone Number Structure"
        TRUE, FALSE /* 대한민국 전화번호 체계에 맞는 수신번호 구조 여부 */
case 5 "Existence of apk"
        TRUE, FALSE /* HTML내의 Header tag 분석을 통해 apk 존재 여부 체크 */
case 6 "Number of word at Smishing message"
        INT /* 스미싱에서 많이 출현하는 단어 수 */
case 7 "Similarity between url domain and related company"
        TRUE, FALSE /* URL, 전화번호와 등록된 회사번호와 비교 */
case 8 "Received Time and Text"
        TRUE, FALSE /* 문자 메시지 수신시간과 내용 비교 */
```

[그림 8] 스미싱 탐지기준 값 세팅

[Fig. 8] The value setting of detection criteria for Smishing

형태소 분류기를 통해 데이터가 만들어지면 스미싱 탐지 기준을 통해 수신을 받은 메시지에 대해 하나씩 값들을 체크하도록 Java를 통해 직접 개발을 하였다. 스미싱 탐지기준에 따른 값을 측정하는 방법은 [그림8]과 같다. 속성은 Message Type, Message Attribute, Number of Emoticons, Phone Number Structure, Existednce of apk, Number of word, similarity of url domain, received time으로 총 8가지의 case를 순차적으로 값을 체크한다.

스미싱 문자 메시지를 학습하기 위한 알고리즘으로 SVM을 사용하였으며, 실험을 위한 툴로는 무료 데이터 마이닝 툴인 WEKA[11]를 사용하였고 버전은 3.7.10이다. WEKA는 데이터 전처리, 분류, 회귀, 클러스터링, 연관 규칙, 시각화 등 데이터 마이닝 툴을 포함하는 오픈소스 프레임워크로서 GUI환경을 제공하여 다양한 데이터 마이닝 알고리즘을 활용할 수 있다. WEKA에서 제공하는 SVM 알고리즘은 SMO(Sequential minimal optimization)으로 SVM 알고리즘 중 가장 널리 쓰이는 알고리즘이다. SMO 알고리즘은 학습단계와 테스트 단계로 학습단계는 학습 데이터를 이용하여 최적의 고차원 평면을 찾게 되고 테스트 데이터를 이용하여 분류를 하게 된다[10].

SMO 알고리즘의 파라미터로는 커널의 종류, 커널 파라미터, C(Complexity parameter), Epsilon 등이 있다. 실험을 위해 SMO 알고리즘의 파라미터의 입력변수인 커널, 커널 파라미터를 변경하며 결과 값의 변화를 도출하였다. 추가로 다른 기계학습 알고리즘과 성능을 비교하기 위해 나이브베이즈(NaiveBayes), 결정트리(Decision Tree)을 같은 데이터를 가지고 학습단계와 테스트단계를 진행하였다.

4.2 실험 결과

본 실험은 SVM 알고리즘 중 SMO를 이용하여 결과를 도출하였다. SMO알고리즘에 Poly Kernel와 RBF Kernel을 이용하여 스미싱 메시지 300개와 일반 메시지 300개, 총 600개의 데이터를 학습시킨 후 랜덤으로 100개의 스미싱 메시지와 일반메시지를 추려내어서 20차례 테스트를 하였다. 입력 값으로는 커널의 종류 및 커널의 종류에 따른 파라미터 값들을 변경시키고, 결과 값으로는 TP(True-Positive) Rate, FP(False-Positive) Rate, Precision, Recall, F-Measure로 값을 측정하였다. 표 3은 스미싱 문자 메시지와 일반 문자 메시지에 대한 실험 결과이다.

[표 3] 스미싱 문자 메시지와 일반 문자 메시지에 대한 SVM 학습 실험 결과
[Table 3] Experimental results of SVM learing between Smishing and ham

입력 값 구분		실험 결과				
커널	커널 파라미터	TP Rate	FP Rate	Precision	Recall	F-Measure
Poly Kernel	Exponent p:1.0	0.931	0.058	0.936	0.931	0.932
Poly Kernel	Exponent p:2.0	0.938	0.012	0.955	0.938	0.942
Poly Kernel	Exponent p:3.0	0.938	0.012	0.955	0.938	0.941
Poly Kernel	Exponent p:4.0	0.922	0.014	0.948	0.928	0.788
RBF Kernel	Gamma g : 0.01	0.875	0.023	0.931	0.875	0.888
RBF Kernel	Gamma g : 0.02	0.907	0.016	0.942	0.923	0.910
RBF Kernel	Gamma g : 0.03	0.894	0.020	0.942	0.894	0.898
RBF Kernel	Gamma g : 0.04	0.875	0.023	0.931	0.875	0.888

커널의 종류와 커널 파라미터 값에 따라서 정확도가 다르게 나온 것을 볼수 있다. 평가 척도로

서 정밀도(precision)과 재현률(Recall)을 통해 계산되는 F-Mearue를 사용하였다. Poly Kernel의 경우 Exponent 변수 p를 2.0으로 설정하였을 때 F-Measure의 값이 0.942로서 가장 높은 정확도를 보였다. RBF Kernel의 경우 Gamma 변수 g를 0.02로 설정하였을 때 F-Measure의 값이 0.910로서 가장 높은 정확률을 보였다.

SVM이 스미싱 분류의 높은 정확도를 보이는 것을 증명하고자 NaiveBayes 알고리즘과 Decision Table 알고리즘을 이용하여 SVM과 같은 환경과 데이터로 실험을 하였다. 결과값은 표 4와 표5에서 볼수 있다. NaiveBayes 알고리즘의 경우 F-Measure의 값이 0.901이고, Decision Table의 경우 F-Measure의 값이 0.883으로 SVM 알고리즘의 정확도보다 현저히 떨어지는 것을 알 수 있다.

[표 4] 스미싱 문자 메시지와 일반 문자 메시지에 대한 NaiveBayes 학습 실험 결과
[Table 4] Experimental results of NaiveBayes learning between Smishing and ham

입력 값 구분	실험 결과				
	TP Rate	FP Rate	Precision	Recall	F-Measure
NaiveBayes	0.891	0.020	0.936	0.891	0.901

[표 5] 스미싱 문자 메시지와 일반 문자 메시지에 대한 Decision Table 학습 실험 결과
[Table 5] Experimental results of Decision Table between Smishing and ham

입력 값 구분	실험 결과				
	TP Rate	FP Rate	Precision	Recall	F-Measure
Decision Table	0.875	0.186	0.899	0.875	0.883

5. 결론

본 논문에서는 문자메시지가 수신되었을 때 스미싱 메시지와 일반 메시지를 분류하는 방법을 제안하였다. 기존에 나와있는 어플리케이션의 경우 content-based 기반으로 탐지를 하고 있기 때문에 단어가 불규칙하게 변경이 될 경우는 오탐률이 높아진다. 본 논문에서 제안하는 방법은 blacklit 기반 기법을 이용하여 일차 분류를 하고, 형태소 분리기에서 전처리과정과 스미싱 메시지에서 빈도수가 높게 나오는 비정상적인 단어들을 저장을 함으로서 명사를 추출하는 정확도를 높임으로서 스미싱 메시지를 탐지하는 오탐률을 줄이는 데 기여를 하였다. 또한, 수신된 문자 메시지들의 특징을 분석하여 스미싱 탐지 기준을 구성하고 SVM 알고리즘, NaiveBayes, Decision Table을 이용하여 스미싱 메시지와 일반문자 메시지를 분류의 정확도를 분석하고 비교하였다. SVM 모델은 스미싱 메시지를 일반 메시지로 분류하는 오류를 최소화하여 NavieBayes, Decision Table과 비교하였을 때 높은 정밀도와 높은 F-measure 결과를 보여주고 있다. 향후 연구로는 SVM 알고리즘의 정확도를 좀더 높일 수 있는 방안으로 SMS의 구조와 SMS 네트워크 통신을 통해 정확도를 높일 수 있는 방안을 연구 분석하고자 한다.

References

[1] Aggarwaly, A., Rajadesingan, A., Kumaraguru, P. "PhishAri: Automatic realtime phishing detection on twitter" In Seventh IEEE APWG eCrime researchers summit (eCRS). Las Croabas, Puerto Rico (2012), 22 – 25.

[2] G. Xiang, J. Hong, C. Rose, and L. Cranor, "Cantina+: A feature-rich machine learning framework for detecting phishing web sites," ACM Transactions on Information and System Security (TISSEC) (2011), vol. 14, no. 2, p. 21.

[3] Y. Zhang, J. Hong, and L. Cranor, "Cantina: a content-based approach to detecting phishing web sites," in Proceedings of the 16th international conference on World Wide Web. ACM (2007), pp. 639-648.

[4] S.J. Delany, M. Buckley, D. Greene, "Sms spam filtering: Methods and data" Expert Systems with Applications, 39 (10) (2012), pp. 9899 – 9908.

[5] YADAV, K., KUMARAGURU, P., GOYAL, A., GUPTA, A., AND NAIK, V. SMSAssassin : Crowdsourcing Driven Mobile-based System for SMS Spam Filtering. In Proceedings of Hot- Mobile 2011: the 12th International Workshop on Mobile Computing Systems and Applications (Phoenix, AZ).

[6] José María Gómez Hidalgo , Guillermo Cajigas Bringas , Enrique Puertas Sánz , Francisco Carrero García, Content based SMS spam filtering, Proceedings of the 2006 ACM symposium on Document engineering (2006), Amsterdam, The Netherlands.

[7] Dae-Neung Sohn , Jung-Tae Lee , Hae-Chang Rim, The contribution of stylistic information to content-based mobile spam filtering, Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, August 04-04 (2009), Suntec, Singapore

[8] http://ko.wikipedia.org/wiki/SVM , Oct 18 (2013).

[9] Hyun woo Kim, Sungyoung Lee, The Phoneme Kernel Technique based on Support Vector Machine for Emotion Classification of Mobile Texts (2013), Journal of KIISE 40(6) 2013.06

[10] Jong Ryul Jin, Dong Seong Kim, Jong Sou Park, The Hardware Design and Implementation of the Support Vector Machines. Proceeding of The Korean Institude of Information Scientists and Engineers, (2004) April 592-594 ; Seoul, Korea

[11] http://www.cs.waikato.ac.nz/ml/weka/ , Oct 18 (2013)

[12] http://www.bok.or.kr, 2013 2Q domestic Internet Banking Service Usage, Oct 18 (2013)

[13] Do-Gil Lee, Sang-Zoo Lee, Hae-Chang Rim, An Efficient Method for Korean Noun Extraction Using Noun Patterns, The Korean Institute of Information Scientists and Engineers (2003), p173-183

[14] Seung-Hyun Yang, Young-Sum Kim, A High-Speed Korean Morphological Analysis Method based on Pre-Analyzed Partial Words, The Korean Institute of Information Scientists and Engineers (2003), p290-301

Authors

이지원 (Ji-Won Lee)

2002년 2월 : 경원대학교 전자계산학과 졸업
2012년 3월 ~ 현재 : 고려대학교 금융보안학과 석사
관심분야 : 정보보호, 스마트폰, 금융보안

이동훈 (Dong-Hoon Lee)

1983년 8월: 고려대학교 경제학과(학사)
1987년 12월: Oklahoma University 전산학 대학원(공학석사)
1992년 5월: Oklahoma University 전산학 대학원(공학박사)
1992년 8월: 단국대학교 전자계산학과 전임강사
1993년 3월~1997년 2월 : 고려대학교 전산학과 조교수
1997년 3월~2001년 2월 : 고려대학교 전산학과 부교수
2001년 2월~현재: 고려대학교 정보보호대학원 교수
관심분야 : 암호프로토콜, 암호이론, USN 이론, 키 교환, 익명성 연구, PET 기술

김인석 (In-Suk Kim)

1980년 2월 : 홍익대학교 전자계산학과 졸업
2002년 2월 : 동국대학교 국제정보대학원 이학석사
2008년 2월 : 고려대학교 정보경영공학전문대학원 공학박사
2011년 3월 ~ 현재 : 국민대학교 컴퓨터공학과 교수
관심분야 : 정보보호, 금융보안