Security Mechanism for Browsers against Spoofing and Phishing Attacks

김정태
목원대학교

스푸핑과 피싱 공격에 강한 브라우저의 보안 메카니즘

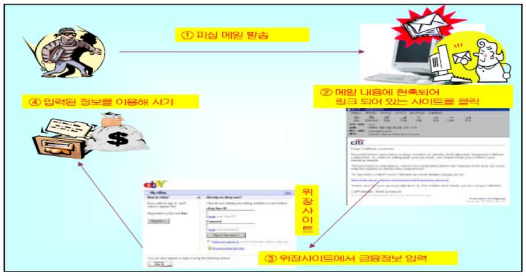
Jung-Tae Kim
Mokwon University
E-mail : jtkim5068@gmail.com

요 약

본 논문에서는 상황기반 개인정보보호 통합 에이전트를 패키지 형태로 공공기관에 공급 공공기관 내부의 개인정보 관리 시스템에 적용되는 피싱 차단 시스템에 대해서 설계하였다. 이러한 시스템은 공공기관 내부 직원의 개인정보 관리 에이전트로 적용과 상황기반 개인정보보호 통합 에이전트로서 금융기관 등의 솔루션에 적용 가능하다. 주요 내용으로는 피싱 유형과 방법을 분석하고 피싱(Phishing) 웹사이트를 탐지, 차단하는 알고리즘을 원천적으로 개발하여 구현하였다.

Ⅰ. 서론

피싱(Phishing)은 1996년 American Online(AOL)을 사용하던 10대들이 일반 사용자들에게 가짜 이메일을 보내는 해킹 기법에서 유래됐으며 이들은 당시, 자신의 이메일을 AOL에서 보낸 이메일이라고 속이는 방법을 통해 일반 사용자들의 계정정보를 훔쳤다. 피싱은 온라인상에서 가짜 미끼를 걸어, 고객의 개인정보(Private Data)를 낚시질(fishing)하는 것을 의미하는 것으로 여기서 Private Data(개인정보)와 Fishing(낚시)의 단어가 합쳐져 피싱이라는 단어가 탄생되었다. 이 단어는 당시 alt.2600이라는 해커가 주로 이용하던 뉴스그룹에서 처음 언급되었다. 피싱에 주로 사용되는 방법은 수신자가 원치 않는 이메일 또는 스팸 등을 발송하여 인터넷 사용자들을 Phisher들이 운영하는 웹사이트로 이동시키는데, 그 웹사이트들은 합법적인 전자 상거래 사이트처럼 위조되어 있다. 여기서 Phisher는 가짜 이메일을 보내어 사용자들을 속이고 정보를 빼내는 사람으로 피싱 공격을 수행하는 사람을 의미한다. 또한 이러한 사이트에서는 사용자에게 계좌 정보를 갱신한다는 명목 하에 패스워드, 주민등록번호, 은행 계좌 혹은 신용카드 번호를 제공하도록 유도한다. Phisher들은 이렇게 획득한 정보를 다양한 용도로 사용하게 되며 피싱 공격의 피해자는 해당 사실을 어느 날 자신이 사용하지 않은 내역이 다수 포함된 고지서 등과 같은 정보를 통해서 알게 될 것이다. 이러한 피싱 범죄의 최근 경향을 다룬 Anti-Phishing Working Group(APWG)의 보고에 의하면 공격이 급증하고 있으며 6월에만 1,422개의 새로운 공격이 APWG에 보고되었으며, 이는 5월에 비해 19% 증가한 수치로 2004년에는 이러한 공격 보고가 매달 52%씩 증가하였다고 한다. (그림1)은 피싱의 기본적인 개념을 보이고 있다 [1].



(그림 1) 피싱의 개념도

II. 국내외 관련 기술 동향

우리나라의 경우 인터넷 뱅킹 등의 금융거래 시 공인인증 및 안전 카드를 사용하도록 되어 있어 피싱을 통해 금융정보가 유출되어도 실제 금융 사기로 연계될 가능성은 외국에 비해 낮은 편이며, 현재는 주로 외국 업체를 사칭하는 피싱이 주를 이루고 있어 영문 메일에 익숙치 않은 우리나라에서는 피해는 적은 편이나 현재는 많은 신고가 이루어지고 있다. 다만, 피싱이 사회적 문제로 제기됨에 따라 국내에서도 모방 범죄가 발생할 가능성이 높으며, 이 경우 정보보호 마인드가 높지 않은 우리나라 이용자들의 피해가 클 것으로 예상되고 있다.

APWG(Anti-Phishing Working Group)의 분석에 의하면 피싱의 위장 홈페이지로 이용된 서버중 우리나라의 서버가 전체의 16%를 차지하여 세계 2위를 기록하고 있다. 우리나라의 경우 인터넷의 발달과 웹호스트 서비스의 활성화로 상대적으로 많은 웹사이트가 구축 운영되고 있으나, 이에 대한 보안 관리 의식은 부족한 편이며, 해킹 등에 쉽게 노출되는 경향이 있다. 보안 취약점을 패치하지 않거나, 서버 설정시 유의해야할 보안 사항을 반영하지 않은 학교, 소규모 비영리단체, 중소기업 등의 웹사이트가 해킹을 당해 피싱에 이용되고 있는 것으로 조사되었다.

<표1> 2007년 피싱 관련 데이터 자료

구분	4 월	5 월	6 월	7 월	8 월	9 월	10 월	11 월
순 URL	55,643	37,438	31,709	30,999	32,079	28,015	34,226	23,630
순 도메인	6,637	5,967	6,006	6,006	5,023	5,058	5,472	5,551
순 URL-도메인 지표	7,622	7,092	7,359	7,538	6,580	6,465	6,704	6,936
순 브랜드	174	149	146	126	129	92	120	178
브랜드 당 URL 수	319.79	251.26	217.18	246.02	248.67	304.51	285.22	132.76

<자료> APWG, 2008 년 1 월

국내의 개인정보보호기술은 아직 시작단계에 있으며, 개인정보보호에 특화된 기술개발보다는 기존 정보보호기술을 개인정보보호에 적용하는 수준에 머무르고 있다. 따라서 클라이언트 기반의 개인방화벽, 서버기반의 방화벽 및 VPN, 클라이언트/서버기반의 암호화기반 기술 등이 개인정보보호를 위한 목적으로 적용되고 있고, 개인정보보호를 위한 전문기술이 개발된 사례로는 개인정보보호정책의 기반 플랫폼 기술인 P3P 기술, 개인정보 검색기술, 개인정보 인증기술, 홈페이지 개인정보 노출 점검 및 차단 기술, 홈페이지 변조를 통한 스파이웨어 삽입 탐지기술 등이 있다. 국내의 대표적인 보안 소프트웨어 메이커인 시만텍은 2006년 12월 26일, 2006년의 피싱 동향을 발표했음. 피싱 공격 건수는 2006년 동안, 계속해서 증가하고 있으며 VoIP나 휴대폰, IM(인스턴트 메시징) 등도 이용된다고 밝히고 있다. 2006년은 피싱사이트, 피싱메일 모두 건수가 증가했음. 겨냥된 것은 대부분이 금융기관이지만 이와 함께 소매, SNS(Social Networking Site), 서비스 프로바이더, 정부기관의 사이트, 인증기관까지 표적이 확대되었다. 현재의 공격 기술로 요일과 계절의 패턴을 보면 여름과 매주 일요일과 월요일은 피싱 메일의 건수가 줄어드는 경향이고 공격측이 회사 측의 휴가와 주말의 쉬는 날을 겨냥하고 있다. 공격 수단은 전자메일에서 사기사이트로 유도한다는 기존형 수법과 함께 전자메일을 사용해 특정 전화번호에 전화를 걸거나 통화료가 저렴한 VoIP 전화로 접촉해 오는 음성 피싱 등도 출현하고 있다. 피싱 공격을 방지하는 기술도 다수 고안되고 있음. 이 중에서 2요소 인증에서는 은행이 이용자에게 하드 디바이스를 배포하고 자주 바뀌는 무작위 번호를 표시하여 계정에 로그온할 때에는 이용자명과 패스워드 이외에도 이 디바이스에 표시된 번호의 입력을 요구하는 구조가 있다. 그러나, 공격측의 기술도 교묘해지고 있으며, 피싱 대응 제품의 대부분은 이미 알려진 피싱사이트의 URL 일람에 따라 사기 사이트를 차단하고 있지만 공격측은 이를 재빨리 빠져나가기 위해 1회밖에 이용할 수 없는 URL을 다수 사용하며 여기서 특정 웹사이트로 연결된 구조를 도입하고 있음. 수 천 개의 URL에서 하나의 사이트로 유도하는 케이스도 있다고 보

고가 되고 있다. 이러한 동향을 총괄하여 시만텍은 금년의 피싱 공격은 빈도가 늘어나 다양해졌으며 '혁신적'으로 되고 있다고 설명하고 있다. 방어측은 항상 이를 상회하는 혁신성이 추구되어 일관적으로 대응책을 강화할 필요가 있다. 따라서, 현재의 국내의 기술동향으로는 본 논문에서 구현하고자 하는 피싱 접속 차단 서비스를 위한 매카니즘에 대한 원천 기술 및 핵심 기술 등이 개발이 되지 않고 있는 실정이다 [2].

III. 지능형 피싱 사이트 접속 차단 기술

피싱(Phishing)은 불특정 다수의 이메일 사용자에게 신용카드나 은행계좌정보에 문제가 발생해 수정이 필요하다는 거짓 이메일을 발송하여 관련 금융 기관의 신용카드 정보나 계좌정보를 등을 빼내는 해킹 기법으로써, 개인정보(Private data)와 낚시(Fishing)의 합성어로 낚시하듯이 개인정보를 몰래 빼내는 것을 말한다. 본 논문에서는 훔쳐가는 피싱의 유형과 방법을 분석하고 피싱(Phishing) 웹사이트를 탐지, 차단하는 알고리즘을 원천적으로 개발하였다.

가. 피싱 공격 방법

피싱의 공격 방법은 매우 다양한 방법이 존재하고 있으며 그중에 몇 가지 방법과 예제를 간단히 설명하면 다음과 같다.

- 유사한 이메일 주소 사용
- 이메일 주소 스푸핑
- 하이퍼 링크위조
- 스크립트를 이용한 주소창 위조
- 팝업창 을 이용한 피싱 방법

나. 피싱(Fishing)의 유형

APWG에 접수된 보고(Phishing Archive)분석을 통해 피싱 사이트의 URL 표현 방법을 크게 세 가지로 분류되고 내용을 소개하면 다음과 같다.

Type 1 : 명시적 피싱 사이트 표현 형태
Type 2 : 유사 도메인명 표현 형태
Type 3 : URL 스푸핑 형태

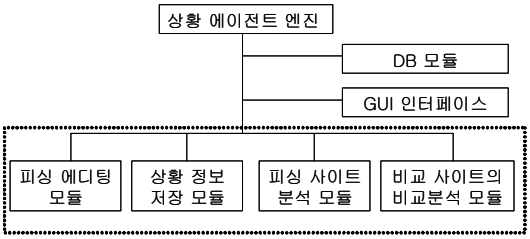
본 논문에서 구현하고자 하는 목표를 구현하기 위하여 사이트 접속 시 피싱 사이트 여부를 점검하는 기능으로 상황인식 기반 지능형 알고리즘을 이용함으로써 피싱 사이트 접속을 통한 개인정보 피해를 방지하는 기능을 수행한다. 피싱 사이트의 탐지 방법은 매우 다양하지만 3가지 형식으로 분리되고 있으며, Post Data가 웹 서버로 전달 될 때 Host IP주소를 정상, 비정상 목록과 비교하여 처리하는 방안을 제안해 보았다. 하지만 정상, 비정상 목록을 전부 데이터베이스화 시키는데 문제가 있어 사용자에게 판단여부 메시지를 주게 되는데 이렇게 되면 사용자에게 불편함을 줄 수 있게 된다. 그 부분은 URL유사도를 측정하는 알고리즘이 추가 되어야 할 것이다. (문자열 유사도 측정) 또한 URL이 완전히 다른 피싱 사이트가 있을 수 있음으로 웹페이지의 소스를 N-Gram과 유사한 알고리즘으로 처리하는 부분을 다음의 모듈로 개발하였다.

- ○ 피싱 에디팅 모듈
- ○ 비교 사이트 비교분석 모듈
- ○ 상황정보 저장모듈
- ○ 피싱사이트 분석모듈

IV. 시스템의 소프트웨어 구조 설계

본 논문에서는 상황인식기반 개인정보보호 통합에이전트를 개발하였으며 시스템은 통합에이전트와 업데이트 서버로 구성되며, 통합에이전트의 모듈 구조를 나타내면 다음과 같다. (그림 2)에서 알 수 있는 바와 같이, 상황인식기반 개인정보보호 통합에이전트는 GUI부, 에이전트 통합관리 모듈의 상위 모듈과 상황인식기반 피싱/파밍 분석모듈, 개인정보 파일관리모듈, 통합정보저장모듈 등 3개의 하위 모듈로 나뉘어진다. 또한 상황인식기반 피싱/파밍 분석모듈은 피싱/파밍 에디팅 모듈, BWG(Black/White/Gray) 사이트 비교분석 모듈, 상황정보 저장모듈, 피싱사이트 분석모듈, 파밍사이트 분석모듈 등 5개의 서브모듈로 구성되며, 개인정보 파일관리 모듈은 파일점검 에디팅모듈, 파일탐색 모듈, 문서필터링 모듈, 프라이버시 필터링 모듈, XML 개인정보 파일관리 모듈 등 5개의 서브모듈로, 통합정보저장모듈은 파일정보 저장모듈, 프라이버시정보 저장모듈,

BWG 사이트 정보 저장모듈 등 3개의 서브 모듈로 구성된다.



(그림 2) 피싱 차단 모듈의 구성도

가. 개발환경

1) 프로그램 사용 환경

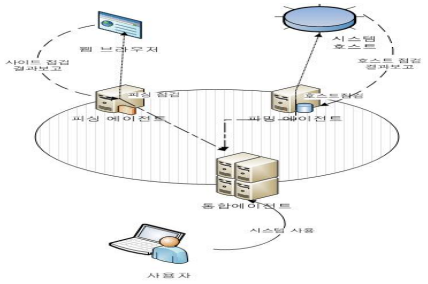
항목	최소사양	권장사양
CPU	Intel Pentium 150MHz	Intel PentiumⅡ 400MHz
RAM	32MB	128MB
HDD	10MB	20MB
VGA	800x600 256Color	1024x768 16bit or 32bit
OS	Microsoft Windows 98/Me/NT/2000/XP (x86, x64)	
Web Browser	Microsoft Internet Explorer 4.0, 5.0, 6.0, 7.0	

2) 프로그램 개발 환경

개발언어	통합 에이전트 시스템	CodeGear Delphi for Microsoft Windows 32
	피싱/파밍 에이전트	
	자동 업데이트	
데이터베이스	피싱/파밍 에이전트	MySQL 4.0.16

나. 시스템 구성도

통합에이전트 시스템의 처리 과정은 3단계로 이루어져있다. 사용자가 사용하려는 프로세스를 검사하는 통합에이전트가 있으며, 각 피싱/파밍/개인정보관리 에이전트는 브라우저나 시스템의 호스트 또한 개인정보를 관리한다.



(그림 3) 시스템 구성도

1) 피싱사이트 분석 모듈

피싱사이트 분석모듈은 상황정보 저장 후 처리되는데, 현재 방문 사이트의 WHOIS 검색을 통한 IP 정보 확인, 화이트 사이트 목록 자동 점검 비교를 통한 내용기반 안전도 점검, 이전 접근 사이트 안전도 분석 등을 통해 종합적으로 판단한다.

2) 파밍 사이트 분석 모듈

파밍사이트 분석모듈도 피싱사이트 분석모듈과 마찬가지로 상황정보 저장 후 처리되는데, 현재 운영 중인 시스템 내의 시스템 폴더의 DNS 변조 확인, 다수 DNS 모의방문결과를 통한 IP 정보 비교 등을 통해 종합적으로 판단한다.

다. 실험 결과

1) 피싱/파밍

현재 운영되고 있는 웹 사이트의 목록을 임의로 (B/W/G)사이트로 등록하여, 상황인식기반 피싱/파밍 에이전트의 실험데이터로 사용하였다.

가) 안전 사이트 실험 데이터

안전사이트 항목은 인터넷 포털(랭키닷컴)에서 제공하는 2008년 8월 사이트 순위 200위중 1~25까지의 항목을 실험데이터로 사용하였다.

나) 차단 사이트 실험 데이터

차단사이트항목은 인터넷(urlblacklist.com)에서 제공하는 2008년 7월 유해 사이트 항목 3천 건의 데이터 중 25가지의 항목을 전제로 실험데이터로 사용하였다.

감사의 글

본 연구는 2008년도 산학연컨소시엄 공동과제의 일부로 수행되었습니다.

V. 결론

본 논문에서는 피싱·파밍 등의 개인정보 침해에 대응하기 위하여 상황인식 기반 피싱·파밍 자동분석 기술을 적용함으로써 인터넷 사이트를 통한 경제활동의 신뢰가 확보되어 인터넷 금융, 온라인 쇼핑몰 등의 인터넷 경제활동을 촉진하게 되며, 개인정보 노출에 따른 피해를 줄일 수 있는 피싱 차단 시스템을 개발하였다.

참고문헌

[1] M. Gray, "Internet Growth and Statistics: Credits and Background" http://www.mit.edu/people/mkgray/net/background.htm

[2 J. Cho and H. Garcia-Molina, "The Evolution of the Web and Implications for an Incremental Crawler," In Proceedings of the 26th International Conference on Very large Databases, Cairo, Egypt, 2000, pp. 5-20.