웹 사이트 보안 위험도 산정 기법

김영갑*, 이준섭**, 조상현***, 김문정*, 이민수**, 김상록**, 김인호****, 김성훈****
*고려대학교 정보경영공학전문대학원
**KAIST 전자전산학과
***엔에이치엔(주) 보안분석팀
****한국정보보호진흥원 기술정책팀
e-mail: *{always, tops}@korea.ac.kr
**{jslee, shcho, mslee, srkim}@dependable.kaist.ac.kr
***bungae@nhncorp.com
****{kih, kimsh}@kisa.or.kr

Method for Evaluating the Security Risk of Website

Young-Gab Kim*, Jun-Sub Lee**, Sanghyun Cho***, Moon Jeong Kim*,
Min-Soo Lee**, Sang-Rok Kim**, In Ho Kim****, Sung Hoon Kim****
*Graduate School of Information Management and Security, Center for
Information Security Technology (CIST), Korea University
**Div. of Computer Science Dept. of EECS,
Korea Advanced Institute of Science and Technology (KAIST)
***NHN Corporation IT Security Analysis Team
****Korea Information Security Agency (KISA)

요 약

최근 전자우편이나 해킹을 통한 피싱과 파밍 등 금융 사기가 많이 발생하고 있다. 이에 이러한 피해로부터 사용자의 경제적 손실 및 개인정보 보호를 위하여 웹 사이트 인증, 전자우편 인증 등의 연구가 진행되고 있다. 기존 인증 방법에서는 WBL (Website Black-List) DB를 사용하였는데, 피싱의 짧은 생명주기(life cycle)로 인해 WBL DB의 유효성은 떨어질 뿐만 아니라, 피싱 사건 발생 후 웹 사이트가 WBL DB에 등록되기 전까지는 확인 불가능하다는 단점을 가지고 있다. 이러한 문제점을 극복하기 위해 WWL (Website White-List) DB를 이용한 연구가 진행 중이지만 아직은 미비한 편이다. 이에 본 논문에서는 기존의 WBL DB와 WWL DB를 이용한 방법이 가지고 있는 한계점을 극복하기 위해 WWL DB 항목을 정의하고, 이를 이용하여 웹사이트 보안 위험도를 정량화할 수 있는 웹사이트 위험도 산정 기법을 제안한다.

1. 서론

최근 전자우편이나 해킹을 통한 피싱(Phishing)과 파밍(Pharming) 등 금융 사기가 빈번하게 발생하고 있다. 피싱은 금융기관이나 쇼핑몰 등을 사칭한 전자우편(E-mail)에 가짜 인터넷 주소를 링크해 금융정보 및 개인 정보를 빼내는 수법이고, 파밍은 피싱보다 한 단계 진화된 수법으로 진짜 사이트 주소를 입력하더라도 가짜 사이트로 접속을 유도해 개인정보를 훔치는 수법을 말한다 [1]. 이에 이러한 피싱, 파밍에 대응하기 위하여 웹 사이트 인증, 전자우편 서버 인증 등에 대한 연구가 많이 진행되고 있다. 기존 웹 사이트 인증 방법에서는 피싱이나 파밍 공격을 데이터베이스(DB)화한 웹사이트 블랙리스트 (Website Black-List, WBL) DB를 사용자에게 제공하는 방식을 사용하였다. 그러나 이러한 WBL 방식의 웹사이트 인증은 몇 가지 단점을 가지고 있다. 즉, 피싱 및 파밍의 생명주기(life cycle)가 짧아서 WBL의 유효성은 떨어질 뿐만 아니라, 피싱 및 파밍 사건 발생 후 웹 사이트가 WBL DB에 등록 또는 신고되기 전까지는 확인 불가능 하다. 이러한 문제점을 극복하기 위해 웹사이트 화이트 리스트 (Website White-List, WWL) DB를 이용한 연구가 진행 중이지만 미비한 편이다. 즉, WWL DB 기반의 인증에서는 단순히 믿을 만한 웹사이트 데이터베이스 구축하고 제공하는 수준에 미치지 못한다. 또한 기존의 WBL DB 기반이나 WWL DB 기반의 인증 방식은 등록된 도메인 정보를 이용하여 웹사이트 (또는 도메인)에 대한 확인만을 수행할 뿐, 웹사이트 위험도를 정량화하거나 웹사이트 내에 존재하는 여러 웹 페이지에 대한 위험 수준을 측정하는 것이 불가능하다. 이에 본 논문에서는 기존의 WBL DB 기반과 WWL DB를 이용한 방법이 가지고 있는 한계점을 극복하기 위해 WWL DB 항목을 정의하고, 이를 이용하여 웹사이트 보안 위험도를 정량화할 수 있는 웹사이트 위험도 산정 기법을 제안한다.

2. 관련 연구

APWG (Anti-Phishing Working Group) [2] 은 피싱,

파밍 공격을 방지하기 위하여 여러 기관 및 기업체 들이 참여하여 활동하는 비영리 조직이다. APWG에서는 참여 기관/업체들의 피싱방지 솔루션을 제공하며 피싱, 파밍, 크라임웨어(Crimeware) 관련 참조 데이터, 연구 데이터, 리소스(Resource) 등 다양한 정보를 제공하고 있다. 또한 국내외 많은 기업체들도 피싱, 파밍 방지를 위한 다양한 솔루션들을 제공하고 있다. 그렇지만 이러한 피싱 방지 기법 및 솔루션들은 WBL DB에 기반으로 하고 있어 앞서 언급한 단점들을 지니고 있다.

또한 앞서 언급하였듯이 피싱, 파밍 공격을 사전에 막기 위해 전자우편 및 웹사이트 인증에 대한 연구가 활발히 진행 중이다. 처음에 전자우편 인증에 대한 연구는 스팸을 차단하기 위한 방법으로 연구가 되어왔으나 많은 피싱 사이트들이 가짜 전자메일을 발송하고, 이를 통해 공격이 이루어짐에 따라 현재 전자메일에 대한 인증은 단순 스팸 차단을 위한 목적을 넘어 피싱을 방지하기 위한 목적으로 연구되고 있다. 전자메일 인증은 크게 마이크로소프트가 중심이 되어 제안한 Send-ID [3] 기법과, Yahoo가 주장한 DomainKey [4] 기법으로 나누어진다. 또한 S/MIME (Security/Multipurpose Internet Mail Extensions) [5] 과 같은 오래 전부터 제안된 방법도 이용한다.

3. 웹 사이트 보안 위험도 산정 기법

이 장에서는 웹 사이트 보안 수준 평가를 위하여 우선 WQL DB의 데이터 항목을 정의하고 이를 이용하여 웹사이트 보안 위험도를 정량화하는 평가 단계를 제시한다.

3.1 웹 사이트 보안 수준 평가 항목 (WWL DB 항목)

본 절에서는 웹 사이트에 대한 보안 수준 평가를 위해 먼저 WWL DB의 세부적인 항목들을 정의하고 그 필요성에 대해 기술한다.

- ServerName : 사이트의 도메인 이름과 그 이름이 등록되어 있는 IP 주소간의 매칭이 정확한지에 대해서 판단한다. 파밍의 경우 도메인과 그 도메인이 등록되어 있는 IP의 정확성을 확인함으로써 방지가 가능하게 된다.

- 도메인국가 : 사용자에 의해 요청된 사이트의 도메인 국가 정보와 실제 배정된 IP의 사용 국가에 대한 정보 확인한다. IP가 할당된 기관의 국가정보와 도메인에 나타난 국가 정보의 일치여부를 확인함으로써 위조된 혹은 중간 경유지를 이용하는 피싱 공격을 판단하는 자료로 사용될 수 있다.

- 도메인수명 : 도메인 수명이란 도메인이 등록된 일부터 도메인 기간이 만료되는 시점까지의 기간을 의미한다. 피싱 사이트 또는 페이지의 경우 잠깐 단기간의 서비스 이후 사라지는 패턴을 고려한 특성이다. 도메인의 수명이 길면 길수록 안전한 특성을 갖는다.

- 도메인 나이 : 도메인 나이란 도메인이 초기 등록된 일자부터 현재까지의 기간을 의미한다. 피싱 사이트는 빠르게 생성되어 단기간에 피싱 행위를 하고 곧바로 사라지는 특성을 갖는 반면, 일반적인 사이트의 경우에는 서비스가 개시되고 일정 기간 지난 뒤에 활발히 이용되는 특성을 보인다. 도메인 나이는 이러한 특성을 반영하기 위한 속성이다.

- 도메인 유명도 : 도메인의 유명도란 검색엔진을 이용한 데이터의 개수를 의미한다. 검색 엔진의 결과 값이 많은 경우, 많은 사용자들이 이용하며 링크를 가지고 있다는 것을 나타낸다. 반대로 검색엔진에 의해 수집된 자료가 없거나 거의 존재하지 않은 경우는 생성된 지 얼마 되지 않은 사이트, 혹은 공개된 시간이 짧은 사이트라고 유추할 수 있다. 따라서 도메인 유명도를 이용한다면 피싱 사이트일 가능성을 좀 더 높일 수 있다.

- DNS 랭킹: DNS 쿼리(Query) 랭킹의 경우 실제 사용자들이 요청한 데이터를 이용한 순위이기 때문에 이러한 특성은 충분히 고려사항이 되어야 한다. 이 방법은 각 사이트에 대한 랭킹 정보를 제공해주는 사이트를 활용하여 이러한 특성의 확인이 가능하다. 이 속성 역시 피싱 사이트의 노출 기간이 짧고 적은 수의 희생자만이 접근한다는 특징을 이용한 것이다.

- 사이트 서비스 : 인터넷 사이트의 서비스 유형에 웹 사이트 위험도가 다르게 평가되어야 한다. 많은 피싱 사이트들은 금융서비스 혹은 쇼핑몰, 커뮤니티 사이트들을 대상으로 피싱을 하기 때문에 이러한 서비스 분류는 위험도 계산에 중요한 속성으로 작용될 수 있다.

- 사이트 구분 : 사이트 구분 속성은 사이트의 서비스를 제공하는 대상이 무엇이냐에 따라 결정된다. 정부, 공공기관과 같이 믿을 수 있는 대상으로부터 서비스가 제공된다면 이 사이트는 위험도가 적을 것이고, 개인에 의해 서비스가 제공되는 사이트라면 앞의 경우보다는 위험도가 높을 것이다.

- 보안 관리자 : 보안 관리자의 유무의 특성은 대상 사이트가 해커에 의해 피싱 사이트로 이용이 될 가능성이 있는지에 대한 판단 기준을 제공한다. 일반적으로 훌륭한 보안 관리자가 있는 사이트인 경우 경유지로 사용되거나 사이트 내부에 피싱 사이트가 존재할 확률은 크게 줄어들 것이다.

3.2 웹 사이트 위험도 수준 측정 단계

본 절에서는 앞서 정의한 위험 수준 평가 항목을 이용하여 웹 사이트의 위험도 측정을 위한 세부 단계에 대해 기술한다. 위험도 측정은 아래와 같이 총 6 단계를 통하여 평가된다.

- 단계1 : 위험측정행렬에 위험요소 정의

본 논문에서는 웹 사이트 위험도를 측정하기 위해 그림 1과 같이 위험측정행렬을 제안한다. 위험측정행렬은 WQL DB에 등록된 항목을 이용할 수 있고, 웹 페이지에 관한 항목을 동시에 이용할 수 있다. 위험측정행렬에서의 위험요소란 위험발생을 야기시키는 요인(조건 또는 상황)을 말하며 앞서 정의된 위험수준 평가항목을 의미한다.

위험점수

위험요소	가중치	4	3	2	1	0	위험지수 (가중치*점수)
…							
…							
…							

(그림 1) 위험측정행렬

- 단계2 : 위험 요소들 사이에 가중치 부여

단계 1에서 정의된 위험 요소들 사이에 가중치(절대적 중요도 또는 상대적 중요도)를 부여한다. 본 논문의 예제에서는 가중치를 5점에서 1점까지 5점 척도를 이용한다.

- 단계3 : 위험 점수 측정

단계 3에서는 각 위험 요소별 위험 점수를 구한다. 위험 점수란 위협발생을 야기시키는 요인이 형성될 수 있는 가능성 정도를 나타내는 것으로 0점 ~ 4점까지 5점 척도를 이용하여 측정한다. 이 때, 위험 점수 4는 위험발생을 야기시킬 가능성이 가장 높음을 의미한다. 각 항목 별 점수 부여는 통계 수치를 반영하여 각각 산출하며 4장의 예제에서 좀 더 자세히 볼 수 있다.

- 단계4 : 전체 위험지수 계산

단계 4에서는 웹사이트의 전체 위험지수(Total Security Index, TSI)를 구한다. 각 위험요소에 대하여 위험점수에 가중치를 곱한 값이 위험지수 (Security Index, SI)라 하고, 개별 위험지수들을 합한 값이 전체 위험지수이다.

- 단계5 : 최대위험지수 계산

단계 5에서는 최대 위험지수(Max Security Index, MSI)를 구한다. 최대위험지수란 웹 사이트가 가질 수 있는 이론적 위험 최대값으로써 본 논문에서는 0점에서 4점까지 측정된 값 중 4점으로 측정된 모든 위험요소들의 위험지수 합을 의미한다.

- 단계6 : 웹 사이트 위험도 계산

마지막으로 웹사이트 위험도(Website Security Risk Index, WSRI)를 구한다. 단계 4와 5에서 구한 SRI와 MSI를 이용하여 WSRI를 다음 식과 같이 정의한다. 보안 위험도 값은 0에서 100 사이의 값을 가지며, 값이 작을수록 안전한 웹사이트를 의미한다.

$$WSRI = TSI / MSI * 100$$

4. 적용 사례

본 장에서는 앞서 제시한 웹사이트 보안 위험도 측정 기법에 대한 이해를 돕기 위해 적용사례를 제시한다.

- 단계1 : 위험측정행렬에 위험요소 정의

본 예제에서는 3.1절에서 정의한 9개의 위험 요소를 이용한다: ServerName, 도메인국가, 도메인수명, 도메인나이, 도메인 유명도, DNS 랭킹, 사이트 서비스, 사이트 구분, 보안관리자

- 단계2 : 위험 요소들 사이에 가중치 부여

정의된 위험 요소들의 가중치는 절대적 중요도 따라 아래와 같은 기준에 의하여 계산된다. 본 예제에서는 3.2절에서 지정하였듯이 5점에서 1점까지 5점 척도로 측정한다.

5 : ServerName, 사이트서비스
4 : 사이트구분
3 : 도메인수명, 도메인나이, 도메인유명도
2 : 도메인국가, DNS 랭킹
1 : 보안관리자

- 단계3 : 위험 점수 측정

각 위험 요소별 위험 지수를 부여하기 위해 그림 2를 이용한다.

항목	설명	4	3	2	1	0
ServerName	등록된 DB값과 값의 일치 여부	불일치				일치
도메인 국가	도메인 국가와 IP 위치 국가 일치여부	불일치				일치
도메인 수명	Expired date - Created Date	2년 미만	2년 이상 4년 미만	4년 이상 6년 미만	6년 이상 8년 미만	8년 이상
도메인 나이	현재시간 - Created date	1년 미만	1년 이상 2년 미만	2년 이상 3년 미만	3년 이상 4년 미만	4년 이상
도메인 유명도	검색 엔진을 통한 도메인 노출 정도	100페이지 미만	100-1000페이지	1000-만 페이지	만-십만 페이지	십만 페이지 이상
DNS 랭킹	랭킹 정보 조회	하위 10%		상위 50%		상위 10%
사이트 서비스 유형	사이트에서 제공하는 서비스의 유형	인터넷 상거래	컨텐츠 제공	포탈서비스 또는 개인웹페이지	뱅킹, 예약, 교육 서비스	
사이트 구분	사이트 운영 주체에 대한 신뢰도	회사	학교	개인	공공기관	정부기관
보안 관리자	보안관리자 능력	보안 관리자 없음		보안 관리자 있음, 보안 관리 계획 없음		보안 관리자 있음, 보안 관리계획 있음

(그림 2) 각 위험 요소의 위험 점수 측정 기준

그림 2의 측정 기준을 적용하기 위해 사용자가 접속하고자 하는 웹사이트를 다음과 같이 가정한다.: 웹사이트의 URL과 등록된 IP 주소가 일치하며, 도메인 국가 정보와

실제 배정된 IP의 사용 국가가 일치한다. 도메인 수명은 5년이고 현재 도메인 나이는 3년이다. 도메인의 유명도는 30여 페이지수 (유명도) 이고, DNS 랭킹은 하우 40%에 속한다. 또한 웹 사이트는 인터넷 전자상거래를 하는 회사다. 웹 사이트 보안 관리자는 존재하나 보안관리 계획 및 규정은 없다.

그림 2 의 측정 기준에 의해 단계 1에서 단계 3까지 진행되었을 때 위험측정행렬은 그림 3 과 같이 작성된다.

위험요소	가중치	4	3	2	1	0	위험지수 (가중치*점수)
Server Name	5					V	0
도메인국가	2					V	0
도메인수명	3			V			6
도메인나이	3			V			6
도메인유명도	3	V					12
DNS 랭킹	2		V				6
사이트서비스	5	V					20
사이트구분	4	V					16
보안관리자능력	1			V			2

(그림 3) 단계 3까지의 위험측정행렬

- 단계4 : 전체 위험지수 계산

제시된 예에서 전체 위험지수(TSI)는 다음과 같다.

TSI = 0+0+6+6+12+6+20+16+2 = 68 이다.

- 단계5 : 최대위험지수 계산

제시된 예에서 최대위험지수 (MSI)는 다음과 같다.

MSI = 4* (5+2+3+3+3+2+5+4+1) = 112

- 단계6 : 웹 사이트 위험도 계산

단계 4와 5에서 구한 SRI와 MSI를 이용하여 WSRI를 다음과 같이 구한다.

WSRI = TSI / MSI * 100 = 68/112 * 100 = 60.71

위 결과로부터 접속하고자 하는 웹사이트는 60.71의 위험을 가지고 있다. 즉, 조금 높은 보안 위험도를 가지고 있어 웹사이트 접속시 주의할 필요가 있다.

앞서 제시한 보안 위험도 산정 기법을 통해 몇몇 웹사이트 보안 위험도를 측정하여 그림 4과 같은 결과를 얻었다. 결과에서 보듯이, NAVER, DAUM, KISA 와 같이 잘 알려지거나 보안관련 기관에서는 낮은 보안 위험도로 측정되어 안전한 웹사이트임을 알 수 있다. 그 외의 사이트들은 보안 위험도가 약간 높거나 아주 높아 의심스러운 사이트이거나 피싱 사이트로 분류되었다.

사이트명	ServerName	도메인 국가	도메인 수명	도메인 나이	도메인 유명도	DNS 랭킹	사이트 서비스 유형	사이트 구분	보안 관리자	TSI	SRI	분류
가중치	5	2	3	3	3	2	5	1	1	100	1	
www.naver.com	0	0	0	0	0	0	2	4	0	14	14,00	안전
			17년	11년	27만개	1위	포털	회사				
www.daum.net	0	0	0	0	0	0	2	4	0	14	14,00	안전
			13년	11년	22만8천개	3위	포털	회사				
www.kisa.or.kr	0	0	0	0	1	0	1	1	0	9	9,00	매우 안전
			11년	11년	5만6천	2273위	공공서비스	공공기관				
www.abc.com	0	0	1	0	4	4	2	4	4	41	41,00	보통
			8년	6년	406개	267,000위	포털	회사				
www.corby.net	0	0	2	2	4	4	2	2	4	48	48,00	보통
			5년	4년	3개	검색불가	개인	개인				
livegame.any.to	0	4	4	4	1	4	4	4	4	71	71,00	위험
			확인불가	확인불가	15100개	검색불가	온라인도박	회사				

(그림 4) 웹사이트 보안 위험도 측정 예시

본 논문에서는 웹사이트 위험도 측정을 위한 항목들을 정의하였지만, 이와 마찬가지로 웹페이지 위험도를 측정하기 위한 항목들(예를 들어, 링크 수, 링크외부도메인 등)을 정의하여 동일한 평가 단계를 통하여 웹페이지 위험도를 측정할 수 있다.

5. 결론 및 향후 연구

본 논문에서는 피싱이나 파밍과 같은 공격에 대하여 웹사이트의 보안 위험도를 산정할 수 있도록 하기 위하여 WWL DB의 데이터 항목을 정의하고 이를 이용한 웹사이트 위험도 산정 기법을 제안하였다. 또한 제안 기법의 이해를 돕기 위해 적용사례를 보였다. 제안한 보안 위험도 산정 기법을 통하여 기존의 WBL DB를 이용한 방식의 단점을 극복할 수 있으며, 웹사이트뿐만 아니라 웹페이지 보안 위험도 측정에도 적용할 수 있다. 향후 WWL DB를 생성 및 업데이트에 대한 연구가 필요하며, 이를 기반으로 적용 가능하도록 웹사이트 보안수준 확인 시스템 구축이 필요하다. 또한, 피싱 방지를 위해서는 이러한 기술적인 접근 외에 피싱 인식 제고 활동, 피싱 정보 공유 및 신속 대응 등의 사회문화적 접근과 피싱 관련 범죄를 처벌할 수 있는 법안 마련 등 법제도적 접근방법이 병행되어야 한다.

참고문헌

[1] Gunter Ollman. “The Phishing Guide - Understanding and Preventing Phishing Attacks.” White Paper, Next Generation Security Software Ltd., 2004.

[2] Anti-Phishing Working Group. http://www.antiphishing.org

[3] Microsfoft, http://www.microsoft.com/mscorp/safety/technologies/senderid/default.mspx

[4] Yahoo, http://antispam.yahoo.com/domainkeys

[5] W. Stallings, "Cryptography and Network Security", Person Education, 2003