Embedded Fingerprinting을 사용한 웹 페이지 생성

박수빈, 조동섭
이화여자대학교 컴퓨터공학과
e-mail: subinio@ewhain.net

Web Page Generation Using the Embedded Fingerprinting

Su-Bin Park, Dong-Sub Cho
*Dept of Computer Science Engineering, Ewha Womans University

요 약

개방형 네트워크 하에서는 상대방이 본인인지 여부와 정보내용이 도중에 변조되지 않았는지 여부 등 안정성 확보가 보장되지 않는다. 그러므로 인터넷상의 전자상거래를 비롯한 네트워크를 통한 사회 경제활동에서도 수기서명, 날인 및 인감등록증명과 같은 기능을 하는 전자서명 및 인증기관에 의한 전자인증이 본격화되기 시작했다.

본 논문에서는 안전한 전자거래를 위한 전자문서 사용에 필수적인 전자서명을 핑거프린팅 기법을 이용하여 웹페이지에 적용하고, HTML로 표현된 웹문서의 안정성을 위한 알고리즘을 제안한다.

1. 서론

정보통신 기술의 발달과 정보통신망의 확충으로 기존의 종이문서가 전자문서로 대체되고 있으며 전자상거래, 계좌이체 등의 금융활동도 증가되는 추세이다. 오프라인으로만 이루어지던 금융 서비스가 온라인상으로 제공됨으로써 거래 활동의 생산성과 거래자의 편의성을 동시에 향상시킬 수 있게 되었다. 그러나 이러한 전자상거래의 장점에도 불구하고 직접 상대방을 확인하고 거래하지 못하는 점에서 다음과 같은 문제가 야기될 수 있다.

인터넷과 같은 온라인상에서는 거래 당사자 간의 신원확인이 어려우며, 디지털 정보의 특성상 전달 과정에서 종이 문서에 비해 위·변조가 상대적으로 용이하다. 즉, 전자문서의 원본과 사본의 구분이 어렵다는 문제점이 있다. 전자문서를 교환하는 과정에서 상대방의 신원과 문서내용의 변조여부가 확인되지 않는 경우에는 사용자 위장의 문제가 있을 수 있으며 스니핑이나 가로채기 공격 등 여러 가지 문제점이 발생할 수도 있다. 따라서 전자상거래에 있어서 전자문서와 관련하여 존재하는 보안상의 위험을 제거하기 위해서는 인증, 무결성, 기밀성, 부인방지의 기능을 가지는 전자서명이라는 안전장치를 필요로 하게 된다. 이는 전자문서, 전자영수증, 대금결제 서명 및 상호인증 등 전자 상거래의 거의 모든 영역에서 주로 사용되는 안전한 상거래와 전자문서 사용에 필수적이 되었다. 또한, 이러한 문제들을 방지하기 위해서 등장한 다른 기술로는 워터마킹, DRM 등이 있는데 본 논문에서는 워터마킹 기술의 한 분야인 핑거프린팅 기법을 이용하여 전자서명을 웹페이지에 적용하여, HTML 문서로 표현된 전자문서의 안정성을 위한 알고리즘을 제안한다.

보 논문의 2장에서는 핑거프린팅과 전자서명 기술의 개요에 대해 기술하고 3장에서는 보다 안전하고 효율적인 HTML 웹문서 핑거프린팅을 위한 방법을 제안한 뒤 마지막으로 4장에서 결론과 향후 연구를 논의한다.

2. 기술 개요

본 장에서는 Embedded Fingerprinting을 사용한 웹 페이지 생성에 필요한 기술인 핑거프린팅과 전자서명에 대해 알아보도록 한다.

2.1 디지털 핑거프린팅

네트워크의 발달과 함께 디지털 이미지나 비디오, 음악, 문서 등 디지털 콘텐츠의 불법적인 복제나 재배포로 인한 지적 재산권 문제가 심각해지기 시작하여 불법적인 복제를 막고 저작권을 효과적으로 보호하기 위한 콘텐츠 보호 기술이 요구되게 되었다.

이러한 요구에 의해 등장한 디지털 워터마킹(Digital Watermarking)은 인간의 의식체계나 감지 능력으로는 검출할 수 없도록 저작권자 또는 판매권자의 정보를 멀티미디어 콘텐츠 내에 삽입하여 추후 발생하게 될 지적 재산권 분쟁에서 정당함을 증명하는데 사용하는 기술이다[1].

(그림 1) 디지털 핑거프린팅의 개요

워터마킹의 한 분야로 디지털 핑거프린팅이 있는데 이 기법은 기밀 정보를 디지털 콘텐츠에 비가시적으로 삽입하는 측면에서는 디지털 워터마킹과 동일하다고 볼 수 있으나 저작권자나 판매자의 정보가 아닌 콘텐츠를 구매한 사용자의 정보를 삽입(그림 1)함으로써 콘텐츠를 불법으로 배포한 자를 추적(trace traitor)할 수 있도록 한다는 점에서 워터마킹과 차별화된 기술이다.

2.2 전자서명

전자서명은 현재 사용되고 있는 서명 또는 기명날인을 전자적 형태로 구현한 것으로서, 전자문서를 작성한 사람의 신원과 전자문서의 변경여부를 확인할 수 있도록 하는 고유정보를 말한다. 이는 광의의 전자서명(Electronic Signature)과 협의의 전자서명(Digital Signature)으로 구분되는데, 광의의 전자서명은 전자펜을 이용한 그래픽 기반의 서명 방식이나 디지털화 서명 등과 같이 서명자 인식만 가능한 전자서명을 말한다. 협의의 전자서명은 비대칭 암호화방식을 이용한 전자서명으로서, 기술적인 면을 중심으로 정의하면, 전자메시지에 해시함수를 적용시켜 메시지요약(message digest)을 만든 후, 이에 공개키 알고리즘과 송신자의 개인키를 이용하여 암호화한 비트의 조합을 말한다. 이러한 전자서명은 서명의 진정성, 무결성, 기밀성 및 부인방지 기능을 확보할 수 있기 때문에 서명으로서의 법적 효력을 인정받고 있다[3].

클라이언트는 서버에게 특정 웹 페이지를 요청한다. 클라이언트는 웹 페이지가 변형 없이 클라이언트의 요청을 받은 서버는 웹 페이지의 내용(html 소스 코드)을 신뢰할 수 있는 특정 소프트웨어를 이용하여 해시 한다. 이렇게 문서를 해시한 결과를 해시 값이라고 하는데 이 해시 값을 서버의 개인키를 이용하여 암호화한 결과가 전자서명이다. 서버는 이 전자서명을 웹페이지에 핑거프린팅 한 후 웹 페이지를 클라이언트에게 보낸다. 클라이언트는 웹 페이지를 받은 다음 내용(웹 페이지의 소스코드)을 소프트웨어를 이용하여 해시한다. (해시 값 A) 그리고 핑거프린팅 된 전자서명의 암호화된 해시 값을 서버의 공개키를 이용하여 복호화 하여 해시 값을 얻는다(해시 값 B). 만약 A와 B가 같다면 클라이언트는 서버가 보낸 내용이 위·변조 없이 원래 그대로 내용이라는 것을 확신할 수 있으며, 만약 공인인증서와 함께 사용됐다면 클라이언트는 내용을 보낸 서버가 인증기관이 인증한 진짜 서버라는 것을 확신할 수 있다.

3. 제안 방식

본 장에서는 HTML 문서로 표현된 웹문서를 기존의 웹문서보다 안전하고 효율적으로 사용하기 위한 알고리즘을 제안한다.

3.1 동작원리

본 논문에서는 핑거프린팅 기법을 웹 문서에 적용시키는 기능을 서버에서 추가하였다. 그리하여 기존의 전자서명으로 주로 이용되던 전자상거래나 인터넷뱅킹 등이 아닌 특정한 웹페이지에 전자서명을 첨부하였다. 이를 통해 웹 페이지 작성자가 전자문서를 작성하였다는 사실과 작성내용이 송·수신과정에서 위조·변조되지 않았다는 사실을 증명한다.

HTML 문서의 소스는 대소문자의 구분이 없고 엔터나 스페이스, 탭은 인정하지 않으며 순차적으로 실행되는 특징을 가지고 있다. 웹 페이지 전자서명에서 핑거프린팅 기법은 원래의 내용에 영향을 미치지 않으면서 제 3의 정보 즉, 암호화된 코드를 입력해야 하기 때문에 핑거프린팅 된 웹 페이지도 원래 웹 페이지와 다르지 않게끔 하여야한다. 위에 언급한 HTML의 이러한 특징들을 이용하면 이진수로 변환한 암호화 코드를 핑거프린팅 하여 여러 가지 방법으로 첨가해 보낼 수 있다.

이러한 과정을 실행하는 곳은 웹 서버인데 웹 서버의 소스는 C나 C++, JAVA 등의 언어로 이루어져 있기 때문에 HTML의 특징과 다르게 대·소문자를 구분하고 엔터나 스페이스, 탭도 인식을 한다. HTML과 다른 이런 특징들을 이용하여 웹 서버에서 읽어 들인 소스 코드에 전자서명을 HTML에서 인식하지 않는 문자를 삽입하여 표현한다. 이를 통해 웹 페이지의 내용에 영향을 주지 않고 핑거프린팅 된 정보는 전달받은 웹 페이지의 소스 속에 감춰져 있어 쉽게 드러나지 않는 방법을 제시한다.

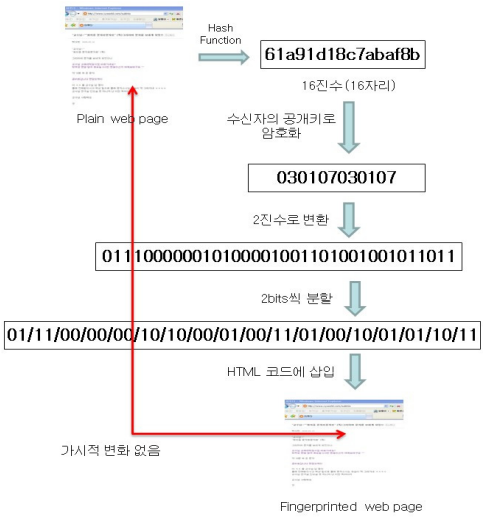
3.2 제안 알고리즘

서버는 사용자가 요청한 웹 페이지에 대해 사용자에게 전송하기 바로 전 단계에서 웹 페이지의 소스코드를 해시하여 message digest를 생성한다. 이 해시된 코드를 이진수로 변환한 후 두 자리씩 나누어서 저장한다. 마지막으로 사용자에게 전송할 웹문서의 HTML 소스 코드에서 제일 처음 읽어 들인 '='부터 순차적으로 2자리씩 삽입하여준다. 본 논문에서는 공백 삽입 시 별다른 영향이 없고, 등장 빈도수가 높은 문자인 '=' 문자를 사용하였는데 이 외에도 다른 문자를 택할 수 있다. 여기서 사용된 공백의 유무에 따른 두 자리의 이진수를 표현하는 코드는 다음에 제시된 <표 1>과 같다.

<표 1> 소스 코드에 따른 표현 코드

소스 코드	표현 코드
A=B	00
A= B	01
A =B	10
A = B	11

다음 코드를 삽입하기 전처리 과정으로 기존의 웹문서에서 '=' 양쪽의 공백을 제거 후 코드에 따른 공백의 삽입을 해주어야한다. 그 후 변환된 이진수를 웹 페이지 소스코드의 맨 위에서부터 '='을 만날 경우 공백을 삽입한다. 이 때 유의할 점은 수식의 '='은 포함시키지 않아야한다는 것이다. '!='나 '==', '>=', '<='와 같은 경우에는 공백문자를 포함시키는 경우 제대로 동작하지 않으므로 위와 같은 문자열을 만날 경우에는 코드 삽입을 수행하지 않고 넘어가는 알고리즘을 추가한다.

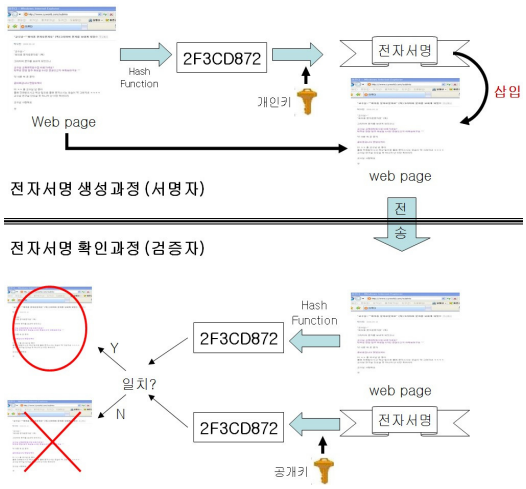


(그림 2) 웹페이지 변환과 삽입

웹 페이지의 변환과 삽입과정은 (그림 2)와 같은데 이를 살펴보면 다음과 같다.

1. 사용자에게 전송할 특정 웹 페이지의 소스코드를 기반으로 해시하여 해시값을 생성한다.
2. 생성된 해시값을 수신자의 공개키를 사용하여 암호화한다. 이 때 양 측은 공개키와 개인키를 이미 가지고 있다고 가정한다.
3. 암호화된 코드를 이진수로 변환한다.
4. 변환된 이진수를 뒤에서부터 2자리씩 끊어 저장한다. 자릿수가 홀수인 경우에는 제일 상위비트에 0을 추가하여 짝수 자릿수를 만들어준다(2bit가 한 쌍이므로).
5. 웹 페이지 소스코드에서 '=' 문자열의 양 옆에 있는 기존의 공백을 제거한다.
6. '=' 문자열의 양 옆에 공백의 유무를 통한 암호화된 코드를 할당한다.

이렇게 변환된 웹 페이지는 최종적으로 사용자에게 전송된다. 사용자는 변환 전 웹 페이지와 같은 화면의 웹 페이지를 받아볼 수 있고 전자서명의 확인을 필요로 할 경우 변환된 코드를 추출하여 복호화한 후 기존의 해시코드와 비교하여 알 수 있다.



(그림 3) 전자서명 생성 및 검증 프로세스

위의 (그림 3)과 같은 전자서명의 프로세스를 이용하여 웹 문서에 적용한 후 검증하는 방법에 대하여 살펴보았다.

4. 결론 및 향후 계획

지금까지 웹 페이지에 전자서명을 첨부하는 방법에 대하여 살펴보았다. 수신자 확인이나 서명의 진정성, 무결성, 기밀성 및 부인방지 기능을 확보할 수 있는 전자서명을 공개키 기반 방식으로 암호화한 후 웹 페이지 안에 공백의 유무에 따른 코드의 할당으로 핑거프린팅 하여 웹 페이지의 내용에 따른 도장과도 같은 역할을 하는 두 가지 기술이 삽입되어 기존의 웹 페이지보다 좀 더 신뢰성 높은 웹 페이지를 받도록 하였다.

기존의 전자서명은 공인인증서를 통해 인증기관에서 인증하는 방식으로 쓰였다면 이 방식은 검증방법을 좀 더 간략화하여 평문의 무결성에 초점을 두어 구현하였다. 웹 페이지의 전자서명은 암호화 시간의 단축과 가정 사항으

로 두었던 키 관리 문제 등 기술적으로 더욱 안정적인 개발이 필요하다.

참고문헌

[1] I. J. Cox, M. L. Miller, and J. A. Bloom, “Digital Watermarking” Morgan Kaufmann Publishers, 2002.
[2] Mohamed Lahcen BenSaad, and Sun XingMing, “Techniques with Statistics for Web Page Watermarking“ PWASET VOLUME 6 JUNE 2005
[3] R. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signatures and public-key crypto systems” Communications of the ACM, vol.21, 1978
[4] 김영준, “전자서명과 인증에 관한 연구”, 통상정보연구, 제3권 제1호, 한국통상 정보학회, 2001.
[5] http://koreainternet.com