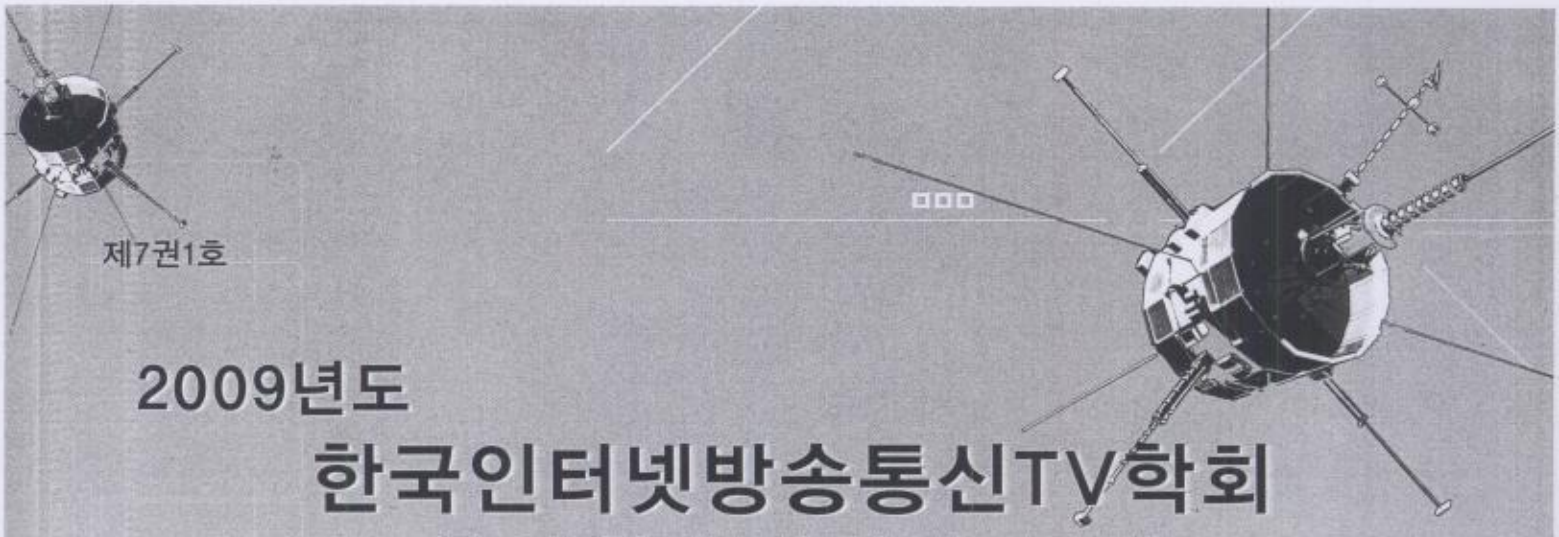

제7권1호

2009년도
한국인터넷방송통신TV학회
춘계학술대회 논문집

일시 : 2009년 5월 29일(금)
장소 : 한국과학기술회관(강남역)
홈페이지 : http://www.iwit.or.kr
주관 및 주최 : (사)한국인터넷방송통신TV학회
(사)인터넷방송통신기술원
동아방송예술대학 성장동력사업단

IWIT (사)한국인터넷방송통신TV학회

논문 목차 (구두)

기본 뉴럴 네트워크를 이용한 데이터 암호화 연구

A Study on Data Encryption using the Basic Neural Network

안성빈, 김영철

Sung-bin Ahn, R. Young-Chul Kim

홍익대학교 컴퓨터정보통신공학과

{ahn, bob}@selab.hongik.ac.kr

요 약

오늘날 인터넷은 개개인의 정보를 전달하는 매개체이며 패킷을 사용해 정보들을 전송한다. 하지만 인터넷으로 전송되는 패킷은 정보가 공개되어 제3자에 의한 패킷 스니핑이 문제이다. 그러므로 전송되는 패킷의 암호화가 필요하다. 그렇지만 기존의 암호화 방법은 일관되지 않은 암호 프로토콜, 사용의 어려움, 암호 어플리케이션의 부재로 인해 암호화를 사용할 수 없는 경우가 많이 있다. 본 논문에서는 이러한 문제를 해결하기 위해서 기본적 뉴럴 네트워크를 이용한 암호화 기법을 제안한다. 이 방법으로 패킷의 정보를 고의적으로 손실시켜 스니핑 공격을 무력화 할 수 있었다. 적용사례로 서버/클라이언트 환경에 적용하여 암호화 기법에서의 사용 가능성을 점검해 보았다.

키워드 : 뉴럴 네트워크(Neural network), 연상 기억장치(Associative Memory), 패킷 스니핑(Packet Sniffing), 암호화(Encryption)

I. 서 론

오늘날 인터넷이 보편화됨에 따라 인터넷은 사람들이 정보를 공유하고 저장하는 수단으로 써 필수적인 역할을 하게 되었다. 이런 중요한 정보들은 패킷의 형식으로 주고받게 되는데, 데이터는 쉽게 공개되어 유출이 가능하다. 패킷을 유출하는 방법에는 여러 가지가 있는데, 본 논문에서는 기존의 패킷 해킹 기법들 중, 패킷 스니핑(sniffing) 기법에 대한 문제 해결을 중점으로 두고 진행하였다.

패킷 스니핑 기법[1]이란, 사용자가 서버와 통신하는 중에 주고받는 패킷에 대한 소통을 공격자가 관찰하다 그 패킷을 가로채는 기법이다. 그리고 이 가로챈 패킷을 분석 하여, 사용자와 서버 간에 어떤 정보를 주고 돌려받는 지, 또 어떤 방식으로 어떻게, 어디서까지 다 알아낼 수 있는 해킹 기법이다.[2]

기존의 패킷 스니핑을 방지하기 위한 방법으로는 네트워크 호스트를 주기적으로 점검하는 방법이 있다. 이러한 점검을 통하여 누가 네트워크를 도청하는지를 탐지하여 조치한다. 다른 방법으로는 스위칭 환경의 네트워크를 구성하여 되도록 스니핑이 어렵도록 한다. 이외에 많은 방법들이 존재하지만 가장 좋은 방법은 데이터를 암호화 하는 것이다. 데이터를 암호화 하게 되면 스니핑을 하더라도 내용을 볼 수 없게 되기 때문에 안전한 전송이 가능한 것이다. 그래서 SSL, PGP[3] 등 인터넷 보안을 위한 많은 암호화 프로토콜이 존재한다.[4]

하지만, 이런 암호화 방법은 일관된 암호 프로토콜의 부재, 사용의 어려움, 암호 어플리케이션의 부재로 인하여 사용할 수 없는 경우가 대부분이다.[5]

위와 같은 문제를 해결하고자 기본적인 뉴럴 네트워크[6]를 이용한 암호화 기법을 제시한다.

기존의 뉴럴 네트워크의 연상기법은 대부분이 일정한 데이터를 복원하거나 선별하는데 이용되어 왔지만[7], 이 논문에서는 패킷을 고의로 손실 시켜 전송하고, 이를 뉴럴 네트워크의 연상 기법을 암호화에 도입하여 새로운 암호화 기법을 제시하고, 적용사례를 통해서 가능성을 검증해 볼 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구로서 기본적인 뉴럴 네트워크의 Associative memory에 대해서 설명한다. 3장에서는 연상기법을 이용한 암호화 방법에 대해서 언급한다. 4장에서는 적용사례로 서버/클라이언트 환경에 적용하여 암호화 기법에 대해 실험한다. 마지막 장에서는 결론 및 향후 연구를 언급한다.

II. 관련 연구

뉴럴 네트워크에서의 Associative Memory (연상기억장치, 결합기억 장치, Content Addressable Memory) 의 개념은 다음과 같다. 일반적으로 대부분의 기억 장치에서는 정보가 저장되어 있는 주소를 입력하여, 입력된 주소에 저장되어 있는 기억 장치의 내용을 접근하게 된다. 하지만 연상 기억 장치에서는 주소를 사용하는 것이 아니라, 접근하려고 하는 자료의 내용을 사용하여 이러한 자료가 저장되어 있는 기억 장치를 접근할 수 있다. 위와 같은 특성에 착안하여 뉴럴 네트워크[8]를 이용한 연상기법은 입력 벡터가 저장된 모든 벡터와 병렬로 비교되면서 일정한 척도로 얼마나 잘 매칭 되는가가 정해지고 가장 잘 매칭 되는 저장 벡터가 선택되어 입력된 벡터를 인식할 수 있다.[9]

연상기법을 활용하기 위해서 수식을 전개하면 다음과 같다.

아래의 식(1)에서 x는 임의의 데이터로 정한다.

$$x = \begin{cases} 1 \rightarrow \text{true} \\ 0 \rightarrow \text{Unknown} \\ -1 \rightarrow \text{false} \end{cases} \tag{1}$$

x는 세 가지의 값을 가질 수 있는데, 1은 항상 참인 값, 0은 알지 못하는 값, -1은 참이 아닌 값을 의미한다.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad X^T = [x_1 \quad x_2 \quad x_3] \tag{2}$$

변형(transformation) 행렬

식(2)에서 X는 x1, x2, x3라는 벡터 값을 가지고 있으며, X를 변형한 변형 행렬 X^T도 x1, x2, x3라는 벡터를 가지고 있다.

위의 식(2)에서의 X와 변형 행렬 X^T, 두 행렬을 이용하여 아래의 수식(3)과 같이 전개를 한다.

$$X \cdot X^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} [x_1 \quad x_2 \quad x_3] = \begin{bmatrix} x_1x_1 & x_1x_2 & x_1x_3 \\ x_2x_1 & x_2x_2 & x_2x_3 \\ x_3x_1 & x_3x_2 & x_3x_3 \end{bmatrix} \tag{3}$$

식(3)과 같이 X와 변형 행렬 X^T의 각각의 벡터들을 곱하여 나온 행렬 데이터를 X_1이라고 한다.

위에서 전개한 식(1), (2), (3)과 같은 방법으로 식을 전개하여 각각의 데이터 X_n을 만들고, 만들어진 데이터들을 아래의 식(4)와 같이 표현한다.

$$\{X_1,\ X_2,\ X_3,\ \cdots\ X_n\} = \mathrm{M} \tag{4}$$

위의 식(4)에서 M은 메모리 내부 상태를 의미한다. 또한 메모리 M은 아래의 식(5)와 같이 전개한다.

$$M = \sum_k X_k X_k^T \tag{5}$$

위의 식(5)와 같이 M은 X와 변형 행렬 X^T들의 곱을 합한 데이터 값으로 정의 되는 것이다.

위와 같은 방법으로 M이라는 Associative memory를 만들고 만들어진 메모리 M을 이용하여 데이터를 손실 시키고 복구 가능하다.

메모리 복구를 위해 아래의 식(6)의 양자화 함수를 이용해야 한다.

$$\phi(x) = \begin{cases} -1 & x<0 \\ 0 & x=0 \\ 1 & x>0 \end{cases} \tag{6}$$

양자화 함수 ϕ 는 X안에 있는 각각의 벡터 값인 x1, x2, x3를 위의 수식과 같이 1또는 0또는 -1로 바꾸어준다.

위의 양자화 함수를 각각의 데이터 X와 메모리 M에 대해서 아래의 수식(7)과 같이 적용을 한다.

$$\phi(X) = [\phi(x_i)]$$
$$\phi(M) = [\phi(x_{ij})] \tag{7}$$

수식(7)을 적용 후에, 손실된 데이터가 y 라고 주어졌을 때, 아래의 수식(8)과 같이 적용하면

$$\therefore\ Z = \phi(\phi(M)y) \tag{8}$$

손실된 데이터 y 로부터 복원 시킨 데이터 Z 를 복구 할 수 있다.

III. 뉴럴 네트워크를 이용한 암호화

인터넷에서 일어나는 대부분의 해킹 기법이 스니핑 기법이

다. 이 스니핑 기법으로 사용자가 사용하는 아이디와 비밀번호, 기타 중요한 정보를 공격자가 중간에서 가로채기를 하여, 사용자에게 피해를 준다. 이러한 스니핑 기법에 대한 대책 방안으로 제시한 연상기법은 연관기억 장치로 위에서 언급한 내용과 같이 사람이 사물에 대한 완전한 기억이 없이 그 사물에 대한 일부의 정보만으로도 어떠한 물체인지 또는 존재인지에 대한 기억을 떠올릴 수 있다. 이런 점을 착안하여, 서버에 접근한 패킷의 완전한 정보를 알지 못한다고 하더라도 연관기억장치를 사용하여, 접근한 정보를 복원 할 수 있다.

즉, 연상기법을 사용하여 사용자가 서버에게 전달하는 패킷의 일부 정보만을 전송한다. 이때 공격자가 패킷을 가로채도 가로챈 정보가 완전한 정보가 아니기 때문에 스니핑 기법의 무력화가 가능하다.

연상기법의 암호화 과정에 대한 알고리즘을 명세하면, 먼저 사용자는 일반적으로 서버에 사용자에 대한 정보를 패킷으로 전송한다. 하지만 이 전송 전에 Associative Memory를 두어 사용자의 패킷에 대한 암호화 하는 과정에서 패킷의 정보를 고의적으로 손실시켜 전송을 한다. 이 손실된 정보를 받는 서버는 손실된 패킷의 대한 정보를 복구할 수 있는 Associative Memory 가지고 패킷 정보를 복구한다. 다시 이와 같은 과정을 거쳐 사용자에게 패킷을 전송한다.

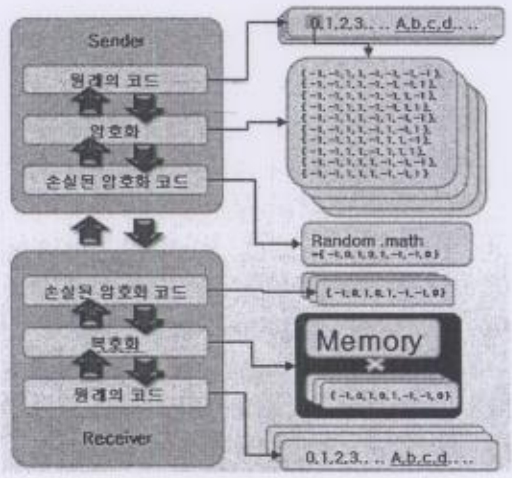

그림 1. 연상기법의 암호화 방법

이러한 과정 중간에 공격자가 패킷 스니핑에 성공하더라도 Associative Memory가 없다면 복호화를 할 수 없다.

그림 1은 사용자와 서버간의 패킷 전송을 연상기법을 적용을 통해서 암호화와 복호화 하는 과정을 간단히 도식화한 모델이다. Sender는 기존의 데이터를 이용하여 전송을 시도하게 되는데, 전송되는 데이터를 연상기법으로 암호화 코드로 만들고, 만들어진 암호화 코드를 랜덤하게 고의적으로 손실시킨다. 손실 시킨 코드를 전송하고 Receiver는 코드를 받은 후에, Associative Memory를 이용하여 손실한 코드를 원래의 코드로 복구 시킨다. 이렇게 원래의 데이터로 복호화 함으로써 안전한 데이터 전송이 가능해지는 것이다.

IV. 적용 사례

적용사례로는 서버/클라이언트 환경 내에서 뉴럴 네트워크를 적용하여 각각의 패킷을 암호화하고 다시 복호화하는 과정을 간단히 모델링 하고 구현을 하였다.

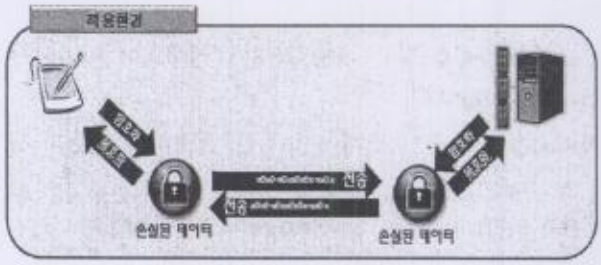

그림 2. 네트워크 전송과정 모델링

그림2는 본 논문에서 설정한 네트워크 환경이다. 사용자는 단일 사용자로 한정을 지었다. 사용자가 패킷 전송 전에 Associative memory를 생성하고 패킷의 값을 임의로 손실을 시켜 그 값을 전송한다. 전송한 패킷은 다시 서버 쪽의 Associative memory가 다시 복구 시키고 그 값을 서버가 받아들이는 방식으로 모델 만들었다.

그림2의 환경을 기반으로 서버/클라이언트 환경에서 단일 사용자로 가정하고 구현을 하였다. 여기선 전송되는 정보는 사용자가 입력한 패스워드라고 가정을 한다.

사용자가 패스워드에 대한 정보에 대해서 위의 그림 3과 같이 입력을 한다. 사용자가 입력한 정보에 대해서 위의 전개한 수식 기반으로 암호화를 한다.

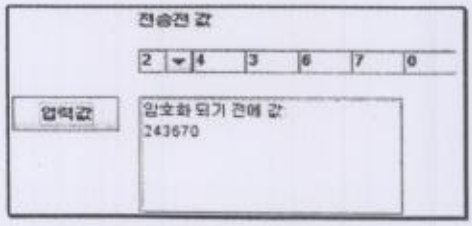

그림 3. 전송 전 값

그림 4는 암호화 시킨 데이터를 나타낸 그림이다.

그림 4. 암호화 시킨 데이터

이 암호화한 데이터를 그림 5와 같이 연상기법으로 산출된 데이터를 랜덤하게 제거하여 손실데이터를 생성한다.

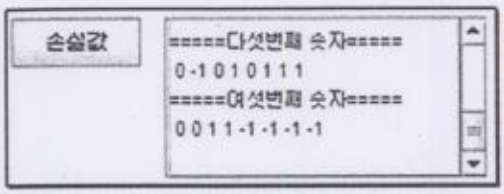

그림 5. Associative Memory를 이용하여 손실 시킨 데이터

손실시킨 데이터를 전송하고, 전송받은 손실데이터를 다시 그림 6과 같이 Associative Memory 이용하여 복구 한다.

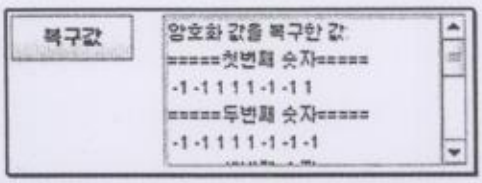

그림 6. Associative Memory를 이용하여 손실시킨 데이터 복구

그리고 복구한 데이터를 그림 7과 같이 원래 사용자가 입력한 데이터로 복호화 한다.

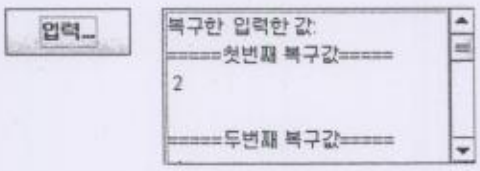

그림 7. 복구한 데이터를 복호화

V. 결론

인터넷 프로토콜은 제3자에 의해 패킷 스니핑이 가능하다. 그래서 전송되는 패킷에 암호가 필요하다. 하지만 기존의 암호화 방법은 일관되지 않은 암호 프로토콜, 사용의 어려움, 암호 어플리케이션의 부재로 인해 암호화를 사용할 수 없는 경우가 많이 있다. 본 논문에서는 이러한 문제를 해결하기 위해 뉴럴 네트워크를 이용한 암호화 방법을 제시하였다. 또한 뉴럴 네트워크의 연상기법을 적용한 암호화 방법을 도구로 구현 하였다. 도구를 이용하여 데이터를 암호화하여 사용 가능성을 점검해 보았다. 그 결과 송신데이터와 수신데이터의 암호화 복호화가 정상적으로 수행되었다. 이로써 공격자의 패킷 스니핑에 대한 안전성 제공과 도구를 사용한 손쉬운 암호화를 할 수 있었다.

향후 연구로 단일사용자로 한정된 모델의 확장을 위해, Associative Memory를 간단한 데이터의 입력의 암호화만이 아닌 다양한 데이터 입력에 대해서 암호화를 제공하며, 쉽게 사용할 수 있는 도구를 개발 중이다. 또한, 패킷 스니핑만이 아닌 다양한 공격자의 공격 패턴에 대한 암호화 방법에 대해서 연구 중이다.

본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터육성지원사업의 연구결과로 수행되었음" (C1090-0903-0004).
본 연구는 교육과학기술부와 한국산업기술재단의 지역혁신인력양성사업으로 수행된 연구결과임.

참 고 문 헌

[1] 최형기 "IPv6 그리고 SCTP 에서의 보안 문제점들", KNOM Review, 2005
[2] 박대우, 윤석현 "VoIP 서비스의 도청 공격과 보안에 관한 연구", 한국컴퓨터정보학회 , 2006
[3] 이준희 "역할 기반 접근 제어에 대한 연구 :웹에서의 RBAC", 홍익대학교 석사학위논문 , 2002
[4] 한승조 "안전성 서비스를 위한 암호화 프로토콜 설계에 관한 연구", 朝鮮大學校生産技術硏究所, 1991
[5] 김용 "패킷 제어 기능을 가진 게이트웨이의 설계 및 구현", 한서대학교 석사학위논문, 2003
[6] 신영숙 "퍼지 인지 맵과 퍼지 연상 메모리를 이용한 오인진단 모델", 한국인지과학회 , 2002
[7] 도양회, 김정우, 배장근 "홀로그래픽 연상 메모리를 이용한 한글 문자 인식", 한국통신학회 , 1993
[8] 김대제, "Neural Network을 이용한 GPS/INS 강결합 항법시스템 구현", 건국대 대학원 석사학위논문, 2004
[9] 이문기, 노육현, 손승임 "Associative Memory를 이용한 패턴 인식 설계", 정보통신 연구 진흥원, 1990. 6