한국ITS학회논문지
제9권, 제2호 (2010년 4월)
pp. 104~109

텔레메틱스 환경에서 무선통신 보안을 위한 사용자 인증에 관한 연구

A Study on User Authentication for Wireless Communication Security in the Telematics Environment

김 형 국*
(Hyoung-Gook Kim)

요 약

본 논문에서는 차량의 무선단말을 이용하여 유선 네트워크와 연결된 정보통신 시스템 서비스를 이용하는 텔레메틱스 환경에서 제3자의 도청이나 공격을 막기 위한 사용자 인증기술을 제안한다. 제안된 사용자 인증방식에서는 차량안의 사용자의 음성신호로부터 생성된 음성 바이오 키를 이용하여 패킷 음성데이터를 암호화하여 정보통신 시스템에 전송하고, 정보통신 시스템 서버에서는 암호화된 패킷 음성데이터로부터 사용자의 음성특징을 복원하여 미리 등록된 사용자의 음성 바이오 키와 비교하여 실시간으로 사용자를 인증한다. 실험을 통해 다양한 공격으로부터 제안된 방식에 대한 안정성을 분석하였다.

Abstract

In this paper, we propose a user authentication technology to protect wiretapping and attacking from others in the telematics environment, which users in vehicle can use internet service in local area network via mobile device. In the proposed user authentication technology, the packet speech data is encrypted by speech-based biometric key, which is generated from the user's speech signal. Thereafter, the encrypted data packet is submitted to the information communication server(ICS). At the ICS, the speech feature of the user is reconstructed from the encrypted data packet and is compared with the preregistered speech-based biometric key for user authentication. Based on implementation of our proposed communication method, we confirm that our proposed method is secure from various attack methods.

Key words: User authentication, telematics, public key infrastructure, gaussian mixture models

Ⅰ. 서 론

최근 자동차에 컴퓨터와 이동통신이 결합되면서 차량에서 운전자의 무선단말을 이용하여 운전자에게 다양한 정보를 실시간으로 제공하는 유선 네트워크의 정보통신시스템 서비스를 이용할 수 있는 텔레메

† 이 논문은 2009년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.
* 주저자 및 교신저자 : 광운대학교 전파공학과 부교수
† 논문접수일 : 2010년 1월 5일
† 논문심사일 : 2010년 3월 10일
† 게재확정일 : 2010년 3월 15일

틱스 환경이 실현되고 있다. 이러한 텔레메틱스 환경에서는 무선 네트워크에서 보안성을 고려한 사용자의 인증의 문제, 보안성과 이동성의 상충성이 선결되어야 한다.

일반적으로 정보시스템에 접근하기 위한 사용자 인증 수단으로 패스워드, PIN 또는 스마트카드 등의 전통적인 방법들이 널리 이용되어 오고 있으나, 최근 들어 지문이나 손 모양, 음성, 홍체, 망막, 혈관, 서명에 이르기까지 개인에 따라 그 특징이 명확하게 다른 신체 부위나 행동 특성으로 사람들의 신원을 확인하는 생체인증 시스템[1]이 연구되고 있으며, 이와 함께 공개키 암호체계(PKI: Public Key Infrastructure) 기술[2]을 이용하여 유무선 네트워크 망의 보안을 강화하는 방식을 사용해 오고 있다.

그러나 이러한 공개암호체계(PKI) 기술은 송신자가 수신자에게 전송하는 정보를 암호화하기 위해서 수신자의 전자인증서에서 추출한 수신자의 공개키로 전송 정보를 암호화하고, 이를 수신한 수신자는 자신만의 알고 있는 수신자 본인의 개인키로 수신한 정보를 복호화하여 그 내용을 확인할 수 있도록 한다. 이때, 공개암호체계(PKI) 기반의 보안 체계를 운용하기 위해서는 누구에게나 공개되는 정보인 공개키가 정말 공개키의 소유주라고 주장하는 사용자의 것인지를 확인할 수 있어야 하며, 이는 신뢰할 수 있는 제3의 기관인 인증기관이 소유주의 공개키에 인증기관의 전자서명을 첨부하여 발행하는 전자인증서를 통해서 확인한다. 그러나 이러한 공개암호체계(PKI) 기반의 보안 체계는 공개키/개인키 쌍의 생성 및 관리, 제3의 인증기관에 의한 전자인증서 발급 및 관리 등의 과정에 따른 시간 및 비용 소모가 큰 문제점이 있다. 또한, 사용자 측면에서 사용자 본인의 개인키 및 공개키 정보를 사용자 단말기 내의 디스크 또는 메모리에 반드시 보관하여 관리해야 하고, 다른 사람의 인증서 정보 역시 보관 및 관리해야 하는데, 이 역시 유지 및 관리에 있어서 비용 소모가 크며 VoIP를 기반으로 하는 통신에 적합하지 않다.

본 논문에서는 차량안의 사용자의 음성신호로부터 생성된 음성 바이오 키를 이용하여 패킷 음성데이터를 암호화하여 정보통신 시스템 서버에 전송하고, 정보통신 시스템 서버에서는 암호화된 패킷 음성데이터로부터 사용자의 음성특징을 복원하여 미리 등록된 사용자의 음성 바이오 키와 비교하여 실시간으로 화자를 인증하는 방식을 제안한다.

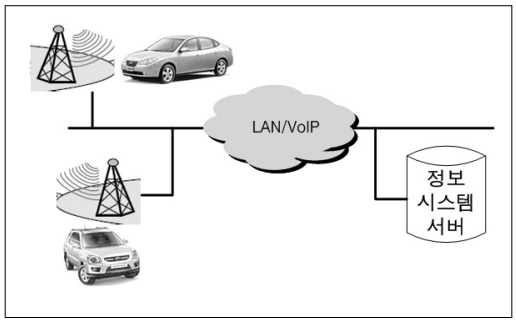
본 논문의 구성은 다음과 같다. II 장에서는 제안된 사용자 인증 시스템에 대해서 소개하며 III 장에서는 제안된 방식을 설명한다. IV 장에서는 구현환경과 실험결과를 소개하고 마지막으로 V 장에서 결론을 맺는다.

Ⅱ. 사용자 인증 시스템의 구성

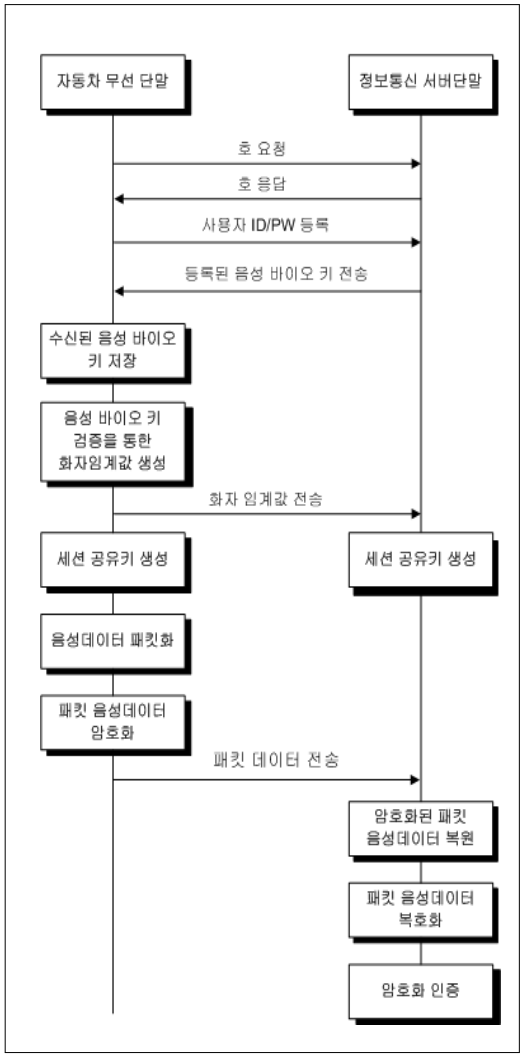
본 논문의 시스템 환경은 <그림 1>에 나타나 있는 바와 같이, 차량에서 무선 네트워크 카드를 장착한 무선단말에서 Wireless Local Area Network (WLAN)과 LAN을 연결해 주는 Access Point(AP)를 통해 WLAN 상의 장치에서 LAN 상의 정보시스템을 이용하는 WLAN 환경을 가정한다.

WLAN 기술을 이용하면 이동전화 기술을 이용할 경우에 속도가 빠르고, 장비의 비용이 10배정도 저렴하기 때문에 무선인터넷 시장에서 경쟁력을 갖추고 있다. 그러나 이러한 WAN은 승인된 사용자에게만 접속을 허용하는 접속에 관한 보안과 스니퍼 등을 이용해 WLAN을 통해 전송되는 내용이 도청되는 취약점을 갖고 있다.

이에 따라 본 논문에서 는 생체정보 중 가장 보편적이며 원격신호처리에 적합한 음성정보를 이용하여 전송채널에서 안전하게 통신하는 방법을 제안한다. <그림 2>는 제안하는 보안통신방법의 통신 과정을



<그림 1> 생체기반 운전자검증 시스템 구조
<Fig. 1> Block diagram of a biometrical driver verification



<그림 2> 무선 보안통신 방법의 흐름도
<Fig. 2> Flowchart of wireless communication security

도시하는 흐름도이다.

우선, 차량안의 사용자가 무선 네트워크를 장착한 단말의 기능 중 VoIP 기반의 통신을 요청함에 따라, 사용자 단말의 제어부는 통신모듈을 제어하여 인바이트(INVITE) 메시지 전송을 통해 WLAN 환경에서 LAN 상의 정보통신 시스템 서버로 호를 요청한다.

이에 대응하여 상대측 정보통신 시스템 서버의 VoIP 통신장치는 호에 대한 응답을 진행한다. 이 과정에서, 상기 VoIP 통신장치들은 통신에 사용할 음성 부호화 및 복호화 방식인 G.711, G.729 방식 등을 설정한다. 이와 함께, VoIP 통신장치가 영상 통신을 지원하는 경우, 영상 신호의 부호화 및 복호화 방식, 예컨대 MPEG4, H.264 방식 등을 설정한다.

시그널링이 수행된 후, 차량안의 무선단말과 정보시스템의 서버가 WLAN을 통해서 연결되면, 정보시스템 서버에 사용자는 자신의 ID와 패스워드를 입력하여 1차적으로 사용자임을 인증받는다.

정보 시스템서버는 1차 검증을 통해 사용자의 ID와 패스워드가 일치하면 사용자가 미리 정보시스템에 등록한 사용자의 음성 바이오키를 서버 저장부로부터 찾아서 사용자의 ID와 음성바이오 키를 Diffie-Helman 알고리즘[3]을 통해 암호화하여 차량의 무선단말에 전송한다.

자동차 무선단말에서는 정보시스템으로부터 수신받은 암호화된 음성 바이오키를 자신의 ID와 조합하여 Diffie-Helman 알고리즘을 통해 음성 바이오키를 해독하여 음성 바이오키를 복원한다.

복원된 음성 바이오키와 차량 안에서 발성된 사용자의 일정구간 음성으로부터 생성된 음성 바이오키를 비교하여 동일한 음성 바이오키 간의 환경변동값을 적용하여 사용자 인증에 필요한 화자임계값을 생성하고 Diffie-Helman 알고리즘을 통해 암호화시킨 후에 정보통신 서버에 전송한다.

그 이후에 미디어 세션이 시작되면 차량의 무선단말 및 연결된 정보시스템 서버에 구비된 세션 공유키 생성부는 각 장치의 메모리에 미리 저장된 음성 바이오키와 새로 수신하여 저장한 제2 바이오 키를 조합하여 Diffie-Helman 수식을 연산하고, 연산 결과값을 이용하여 세션 공유키를 생성한다.

세션 공유키의 생성이 완료된 후, 차량안의 무선단말이 SRTP 세션을 시작한다. 즉, 차량안의 무선단말에 구비된 외부 입력장치로부터 영상 또는 음성 데이터가 입력되면, 부호화/복호화 모듈로 전달되고, 부호화/복호화 제어부는 호 설정 과정에서 선택된 부호화 방식에 대응하는 하나의 부호화부를 선택한다. 이에 따라, 외부 입력장치로부터 영상 또는 음성 데이터는 부호화부를 통해 부호화된 후, 통신모듈로 전달된다. 이에 대응하여, 통신모듈은 부호화된 영상 또는 음성 데이터를 패킷화한다.

다음으로, 통신 모듈은 패킷화된 데이터 상기 세션 공유키를 이용하여 암호화한 후, 정보시스템 서버로 암호화된 패킷 데이터를 전송한다.

정보통신 시스템 서버는 통신모듈을 통해 암호화된 패킷 데이터를 수신하고, 상기 세션 공유키를 이용하여 패킷 데이터를 복원한다.

복원된 데이터는 부호화/복호화 모듈로 전달되고, 부호화/복호화 제어부는 호 설정 과정에서 선택된 부호화 방식에 대응하는 하나의 복호화부를 선택하고, 통신모듈로부터 입력되는 부호화된 영상 또는 음성 데이터는 복호화부를 통해 복호화된다.

복호화된 영상 또는 음성 데이터는 외부 출력장치로 전달되기 이전에, 암호화 인증을 수행하게 된다. 즉, 복원된 패킷 데이터로부터 사용자 음성특징을 추출하고 메모리에 저장된다.

암호화 인증부는 세션 공유키를 생성하는 과정에서 수신한 음성 바이오 키와 사용자 음성특징을 서로 비교하여 동일성 여부를 확인한다. 나아가 암호화 인증부는 음성 바이오키와 수신한 음성신호로부터 추출한 음성특징이 서로 동일한 경우에 한하여 암호화 인증이 성공적으로 수행된 것으로 확인한다.

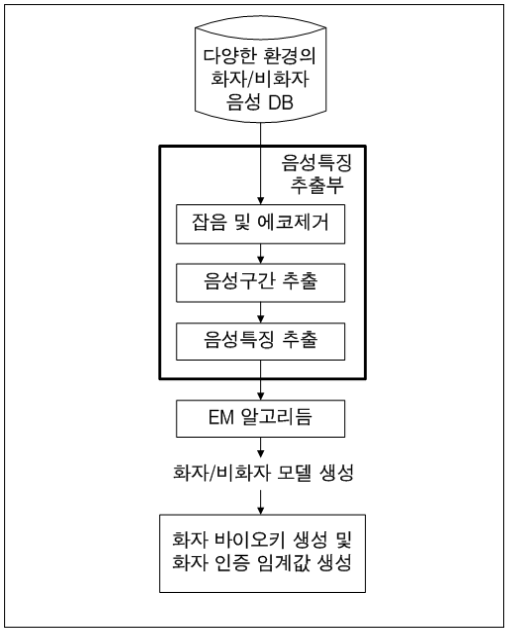
마지막으로, 암호화 인증이 성공적으로 수행되면, 복호화된 상기 영상 또는 음성 데이터를 외부 출력장치로 전달하여 출력한다.

위에 전술한 바와 같이, 제안된 VoIP을 기반으로 하는 통신 방법 및 장치는 공개키 기반구조의 구축 여부에 상관없이 단말에서 미디어 데이터에 대한 암호화를 용이하게 실시할 수 있다. 또한, 미디어 세션에서 키 교환 및 인증이 수행되므로 다양한 시그널링 프로토콜에 적용할 수 있으며, 사용자의 생체 정보를 이용하여 음성 디지털 키(음성 바이오키)를 생성하므로, 통화의 보안을 유지할 수 있다.

Ⅲ. 화자인증 알고리즘

본 논문에서 제안한 화자인식 알고리즘은 크게 음성 바이오키 생성과 화자음성 인증으로 구성되고, 음성 바이오키 생성의 흐름도는 <그림 3>과 같다.

먼저, 전화통화를 통해서 녹음된 다양한 환경에서



<그림 3> 음성기반 바이오키 생성 알고리즘 구조
<Fig. 3> Block diagram of a speech-based biometric key generation algorithm

의 화자 및 비화자들에 대한 음성집합들로부터 잡음과 에코 성분를 제거한다. 본 논문에서는 [4]에서 사용한 음성향상 알고리즘을 사용하여 잡음과 에코를 제거하는 방식을 사용한다. 즉, 첫 단계로 크로스 스펙트럼 추정에 기반한 적응필터를 통해 에코를 제거하고, 두 번째 단계로 Log Spectral Amplitude 음성추정 방식 추정을 통해 외부 배경잡음을 제거하여 음성의 음질을 향상시킨다.

다음 단계로 잡음과 음성이 제거된 음성신호로부터 각 프레임의 절대 에너지가 설정된 값보다 큰 구간을 실제 음성구간으로 추출하여 메모리에 저장하고, Mel-scale Frequency Cepstral Coefficients (MFCC) 추출 방식[5]을 사용하여 음성의 특징 정보를 추출한다. 이 절차는 pre-emphasis 처리, 해밍 윈도우 처리, fast Fourier transform을 각각 수행하여 주파수영역으로 변환된 값들을 Mel 대역 필터뱅크를 통과시키고, 로그화, discrete cosine transform을 거쳐 식(1)을 통해 프레임당 13개의 MFCC 계수값을 획득한다.

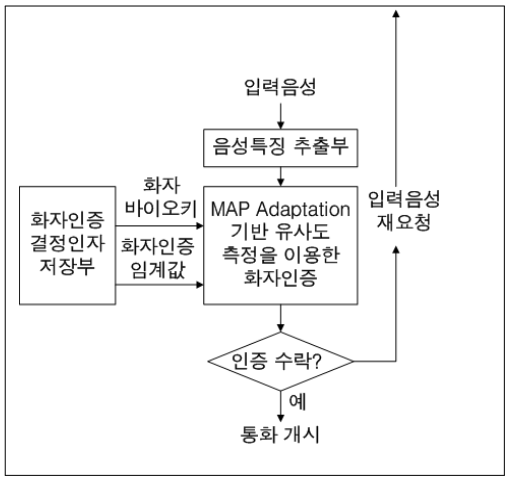
$$X(n)=\sum_{k=1}^{M}\log(S(k))\cos\left[(k-0.5)\frac{\pi}{M}\right],\quad n=1,\ldots,L \tag{1}$$

여기서, S(k)는 k번째 Mel-filter bank를 통과한 신호이고, M은 필터뱅크 개수, L은 MFCC 차수이다.

추출된 음성특징들은 여러 개의 가우시안 확률밀도(Gaussian probability density) 함수들에 각각의 가중치를 준 다음, 이를 선형 결합함으로써 임의의 모양을 갖는 확률밀도 함수를 표현하는 Gaussian Mixture Model(GMM)의 Expectation and Maximization(EM) 학습알고리즘[6]을 통해 화자 및 비화자 모델을 생성한다. 생성된 화자 모델의 평균값을 이진화하여 음성 바이오키를 생성하고, 이와 함께 화자와 비화자를 식별할 수 있는 화자 임계값을 생성한다. 즉, 음성의 특징 파라미터 벡터의 확률분포는 화자마다 그 모양이 다르며, 이러한 확률분포를 GMM을 이용하여 모델링하여 인증하고자 하는 화자의 모델로 사용함으로써 화자인증에 적용한다.

<그림 4>는 본 논문에서 사용한 GMM 기반의 화자인증 알고리즘 구성도를 나타낸다.

화자인증은 발성된 음성이 원하는 화자 즉 의뢰인인지 또는 사칭자인지를 구분해 내는 것으로 의뢰인에 대한 초기등록이 요구된다. 시스템을 사용할 의뢰인들에 대한 화자모델이 사전에 모두 만들어지고 또한 각각의 의뢰인에 대한 배경화자들을 코호트(cohort) 방법으로 선정을 한다. 다음으로 화자인증 시스템에 입력으로 들어온 음성 데이터(X)를 사용하여 다음의 식과 같이 유사도(likelihood ratio) Λ(X)를 계산한다.



<그림 4> GMM 기반의 화자인증 알고리즘 구조
<Fig. 4> Block diagram of a GMM-based speaker authentication algorithm

$$\Lambda(X)=\log p(X|\lambda)-p(X|\lambda_B) \tag{2}$$

여기서 λ는 의뢰인의 화자모델이고, λ_B는 의뢰인의 배경화자모델을 나타낸다.

유사도에 사용하는 배경 화자수와 배경화자 선택 방법은 Maximal Spread Close 방식과 Maximal Spread Far 방식을 혼용한 방식을 사용하였다.

최종적으로, 계산된 유사도 값인 Λ(X)을 정해진 문턱값 Θ와 비교하여 다음과 같이 인증을 수락 또는 거부한다[5].

$$\begin{aligned}&\text{화자인증: } \Lambda(X) > \Theta\\ &\text{화자거부: } \Lambda(X) < \Theta\end{aligned} \tag{3}$$

Ⅳ. 실험 및 결과고찰

본 논문에서는 시스템 구현을 위해 사용자 단말과 화자인증 서버는 Window XP 환경에서 구현하였고, 음성정보처리 관련 모듈, 보안모듈(Security: SHA-256, AES, RSA), 네트워크 패킷 송수신 모듈, 프로토콜 처리 모듈을 구성하여 실험을 수행하였다.

화자인증에 사용한 음성데이터는 마이크를 통해 8kHz/8 비트의 PCM 방식으로 녹음하였으며 끝점검출과 전처리 과정을 거친 후, 20 ms의 프레임 크기와 10ms의 프레임 간격을 사용하여 13차 MFCC 계수, 13차 delta-cepstral 계수를 포함하여 총 26차의 음성특징을 추출하였다. 화자 및 비화자 GMM 모델링은 가우시안 믹쳐의 개수를 16 개로 고정하여 사용하였다.

화자검증을 위해서는 차량 안에서 발성한 화자 20명이 각각 30문장을 발성한 음성 DB를 구성하여 문장독립형 화자인증 실험에 적용하여 89.5%의 화자인증 정확도를 획득하였다.

보안성 평가를 위해서는 오프라인 사전 공격, 전방향 공격, Denning-Sacco 공격, 능동적 중간자 공격을 총 100번을 수행하여 안전함을 확인할 수 있었다.

Ⅴ. 결 론

본 논문에서는 차량의 무선단말을 이용하여 유선 네트워크와 연결된 정보통신 시스템 서비스를 이용하는 텔레메틱스 환경에서 제3자의 도청이나 공격을 막기 위해 음성 바이오키를 이용하여 사용자를 인증하는 기술을 제안하였다. 제안된 방식의 강인성은 실제 텔레메틱스 환경에서 매우 유용하게 사용될 수 있으리라 생각된다.

향후 연구과제로는 휴대폰과 같은 단말기에 화자인증시스템을 구축하여 저전력과 가용 메모리가 적은 환경에서 유용하게 사용할 수 있는 화자인증 방법에 대한 연구를 수행하고자 한다.

참 고 문 헌

[1] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4-20, Jan. 2004.

[2] C. Adams and S. Lloyd, *Understanding Public-Key Infrastructure: Concepts, Standards, and Deployment Considerations*, Macmilan Technical Publishing, 1999.

[3] J. Fry and M. Langhammer, “RSA & Public Key Cryptography in FPGAs,” *Tech. Rep*., Altera Corp., 2005.

[4] H. G. Kim, “"Implementation of chip and algorithm of a speech enhancement for an automatic speech recognition applied to telematics device,” *한국ITS학회논문지*, 제7권, 제5호, pp. 90-96, 2008. 10.

[5] S. Molau, M. Pitz, R. Schluter, and H. Ney, “Computing mel-frequency cepstral coefficients on the power spectrum,” *Proc. ICASSP,* vol. 1, pp. 73-76, May 2001.

[6] D. A. Raynolds, T. F. Quatieri, and R. B. Dunn, “Speaker verification using adapted Gausssian mixture models,” *Digital Signal Processing*, vol. 10, pp. 19-41, Jan. 2000.

저자소개

김 형 국 (Kim, Hyoung-Gook)

2007년 3월 ~ 현재 : 광운대학교 전파공학과 부교수