Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.26, NO.1, Feb. 2016 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2016.26.1.69

HV-KEM을 이용한 생체 정보 기반 인증 프로토콜*

서 민 혜,[1†] 황 정 연,[2] 김 수 형,[2] 박 종 환[3‡]
[1]고려대학교, [2]한국전자통신연구원, [3]상명대학교

Biometric Authentication Protocol Using Hidden Vector Key Encapsulation Mechanism*

Minhye Seo,[1†] Jung Yeon Hwang,[2] Soo-hyung Kim,[2] Jong Hwan Park[3‡]
[1]Graduate School of Information Security, Korea University,
[2]Electronics and Telecommunications Research Institute,
[3]Department of Computer Science, Sangmyung University

요 약

생체 정보를 이용한 사용자 인증은 사용자가 어떠한 정보도 소유하거나 기억할 필요가 없으므로 사용자 측면에서 매우 편리하다. 그러나 생체 정보는 한번 노출되면 영구적으로 사용할 수 없기 때문에 인증 수행 시 생체 정보의 프라이버시 보호가 필수적이다. 또한 생체 정보는 본질적으로 노이즈가 있기 때문에 인증 수행 시 적절한 오차 범위 이내에서 이를 처리할 수 있어야 한다. 최근 퍼지 추출기(fuzzy extractor)를 이용한 생체 정보 기반 인증 프로토콜에 대한 연구가 활발히 진행되고 있으나, 사용자가 헬퍼 데이터(helper data)를 추가적으로 기억해야 하는 문제점이 존재한다. 본 논문에서는 함수 암호 중 하나인 HV-KEM(Hidden Vector Key Encapsulation Mechanism)을 이용하여 사용자가 어떠한 정보도 소유하거나 기억할 필요 없는 생체 정보 기반 인증 프로토콜을 제안한다. 또한, 인증 프로토콜 설계를 위한 HV-KEM의 보안 요구사항을 정립하고 제안하는 인증 프로토콜을 정확성/안전성/효율성 측면에서 분석한다.

ABSTRACT

Biometric authentication is considered as being an efficient authentication method, since a user is not required to possess or memorize any other information other than biometrics. However, since biometric information is sensitive and could be permanently unavailable in case of revealing that information just once, it is essential to preserve privacy of biometrics. In addition, since noise is inherent in the user of biometric recognition technologies, the biometric authentication needs to handle the noise. Recently, biometric authentication protocols using fuzzy extractor have been actively researched, but the fuzzy extractor-based authentication has a problem that a user should memorize an additional information, called helper data, to deal with their noisy biometric information. In this paper, we propose a novel biometric authentication protocol using Hidden Vector Key Encapsulation Mechanism(HV-KEM) which is one of functional encryption schemes. A primary advantage of our protocol is that a user does not need to possess or memorize any additional information. We propose security requirements of HV-KEM necessary for constructing biometric authentication protocols, and analyze our proposed protocol in terms of correctness, security, and efficiency.

Keywords: Hidden Vector Key Encapsulation Mechanism, Biometric Authentication

Received(11. 19. 2015), Modified(01. 05. 2016), Accepted(01. 11. 2016)

* 이 논문은 2016년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.B0126-15-1007, 상황인지 기반 멀티팩터 인증 및 전자서명을 제공하는 범용 인증 플랫폼기술 개발)

† 주저자, smh89122@korea.ac.kr
‡ 교신저자, jhpark@smu.ac.kr(Corresponding author)

I. 서 론

1.1 개요

IT의 기술의 발전으로 스마트폰과 같이 일상에서 사용하는 디바이스를 이용하여 언제든지 인터넷에 접속할 수 있으며, 다양한 네트워크 서비스들이 제공되고 있다. 이에 따라 다수의 개인 정보들이 온라인상에서 취급되고 있으며, 개인 정보에 대한 보안이 중요한 이슈로 떠올랐다. 정보의 유출을 방지하기 위하여 현재 대부분의 서비스들이 사용자 인증을 통해 개인 정보에 대한 접근 권한을 관리하고 있다. 대표적인 사용자 인증 방법으로는 아이디/패스워드 기반 인증[1], 인증서(certificate) 기반 인증[2], OTP (one-time password) 기반 인증[3]이 있다.

OTP 기반 인증 방법과 인증서 기반 인증 방법은 각각 사용자가 OTP 기기와 인증서를 소유하고 있어야 하며, 패스워드 기반 인증 방법은 사용자가 자신의 아이디에 대한 패스워드를 기억하고 있어야 한다. 이처럼 기존의 인증 방법들은 사용자가 소유하거나 기억하고 있는 정보를 이용하여 인증을 수행한다. 하지만 생체 정보 기반 인증 방법은 사용자가 별도로 소유하거나 기억하고 있는 정보가 필요하지 않고, 지문이나 홍채와 같은 생체 정보를 이용하여 인증을 수행한다. 따라서 사용자 입장에서 사용이 편리하며, 기존의 인증 방법들에 비해 높은 본인 상관성을 가진다. 최근에 생체 정보 기반 인증을 이용한 다양한 서비스가 등장하면서 이에 대한 관심이 더욱 높아지고 있다[4, 5, 6].

생체 정보를 기반으로 하는 인증 방법은 다음과 같은 요구사항을 만족해야 한다. 먼저, 노이즈를 처리할 수 있어야 한다. 생체 정보는 기본적으로 노이즈가 많은 데이터이기 때문에, 동일한 사용자에 대해서도 생체 정보를 측정할 때마다 조금씩 오차가 발생한다. 따라서 사용자 인증을 위하여 생체 정보의 허용 가능한 오차범위를 설정하고, 오차범위 내에 속하는 데이터에 대해서 인증 수행이 가능해야 한다. 또한, 인증 수행 과정에서 생체 정보의 프라이버시를 보호해야 한다. 생체 정보는 사용자 고유의 정보로 사용자가 임의로 변경할 수 없으며, 한번 노출되면 영구적으로 사용이 불가능하게 된다. 따라서 지속적인 인증 수행을 위하여 생체 정보에 대한 프라이버시가 보호되어야 한다.

기존의 생체 정보를 이용한 인증 기술은 노이즈가 많은 생체 정보로부터 하나의 키를 추출해내는 알고리즘에 근거하여 설계되었다. 추출되는 키가 암호학적 프리미티브에 사용되기 위해서는 충분한 엔트로피(entropy)를 보장해야 하며, 이를 위해 퍼지 추출기(fuzzy extractor)라는 개념이 등장하였다[7]. 이후 퍼지 추출기를 이용한 생체 정보 기반 인증 기술에 대한 연구가 활발히 진행되었다[8, 9, 10]. 퍼지 추출기를 이용한 인증 기술은 생체 정보의 노이즈를 처리할 수 있고 생체 정보에 대한 프라이버시를 보장하지만, 사용자가 헬퍼 데이터(helper data)라는 추가 정보를 기억하고 있어야 하는 불편함이 존재한다.

1.2 기여도

본 논문에서는 생체 정보를 이용한 새로운 인증 프로토콜을 제안한다. 새로운 인증 기법은 최근에 연구되는 함수 암호(functional encryption)[11, 12] 중 하나인 HV-KEM(Hidden Vector Key Encapsulation Mechanism)을 이용하여 설계된다. 제안하는 인증 프로토콜의 장점으로는, 첫째, 생체 정보의 측정 시 발생할 수 있는 오차를 일정 확률 이하로 허용하면서 사용자 인증을 처리할 수 있다는 것이다. 즉, 등록 단계에서 측정된 생체 정보와 인증 단계에서 측정된 생체 정보 간에 일정 확률 이하로 오차가 발생하면 인증이 성공하도록 구성할 수 있다는 것이다. 둘째, 사용자의 생체 정보는 토큰과 암호문 형태로 암호화되어 제3자에게 노출되지만, 그 값들을 통해서는 사용자의 생체 정보를 얻는 것이 어렵다는 것이다. 즉 생체 정보의 프라이버시를 보장할 수 있다. 셋째, 사용자는 인증에 필요한 어떠한 정보를 기억하거나 저장할 필요가 없다는 것이다. 초기 생체 정보를 이용한 등록을 마치면, 이후 인증이 필요할 때마다 사용자는 인증서버에 인증을 요청하기만 하면 된다. 즉, 사용자의 기기에 신뢰영역(예를 들어, SE(Secure Elememt)나 trust zone 등)을 갖출 필요가 없다. 넷째, 생체 정보를 측정하는 인지(recognition) 알고리즘은 난독화(obfuscation) 등 암호학적으로 필요한 비밀키 삽입 등의 가정이 필요 없으며, 생체 정보 측정 시 발생하는 에러를 보정하기 위한 추가적인 값을 안전하게 저장할 필요도 전혀 없다. 단순히 공개된 (그리고 오차확률이 적은) 인지 알고리즘을 활용하여 생체 정보를 추출하면 된다.

또한 본 논문에서는 위와 같은 장점을 갖는 생체 정보 기반 인증 프로토콜이 HV-KEM으로부터 설계될 수 있음을 보인다. 이를 위해 HV-KEM에 필요한 안전성을 두 가지 - 속성 정보 프라이버시와 암호문 위조불가능성 - 개념으로 정립하고, 각각의 안전성에 필요한 공격 모델을 제안한다. 그리고 기본적으로 동일성 술어(equality predicate) 기능을 제공하는 HV-KEM을 이용하여 어떻게 생체 정보의 측정 오차를 처리하는 기법으로 변환할 수 있는지에 대한 아이디어를 제시한다. 현재까지 HV-KEM을 이용하여 생체 정보 기반의 인증 프로토콜에 제시한 사례는 없는 것으로 보인다. 그 대신 HV-KEM이 사용될 수 있는 예로는 의료 정보를 환자의 속성으로 분류하여 암호화하고, 나중에 환자의 프라이버시를 보호하면서 특정 환자에 대한 의료기록만 검색하는 의료 정보 검색암호가 제시되었다[12].

본 논문의 구성은 다음과 같다. 제2장에서는 HV-KEM 기법의 정의와 보안 요구사항에 대해서 설명한다. 제3장에서는 HV-KEM을 이용한 생체 정보 기반 인증 프로토콜을 제안하고, 제4장에서는 제안하는 인증 프로토콜을 정확성/안전성/효율성 측면에서 분석한다. 마지막으로 제5장에서 결론을 맺는다.

II. HV-KEM(Hidden-Vector Key Encapsulation Mechanism) 기법

2.1 알고리즘

HV-KEM 기법은 다음과 같은 네 개의 알고리즘으로 구성된다[12].

- **설정(Setup)** (k,n) : 보안 상수(security parameter) k와 속성 벡터의 차원 n을 입력으로 받아 공개키(PK)와 비밀키(SK)를 생성하는 알고리즘이다.
- **토큰생성(GenToken)** $(SK, \vec{\sigma})$: 비밀키 SK와 n차원 속성 벡터 $\vec{\sigma}$를 입력으로 받아 $\vec{\sigma}$에 해당하는 토큰($TK_{\vec{\sigma}}$)을 생성하는 알고리즘이다.
- **캡슐화(Encap)** $(PK, \vec{x})$: 공개키 PK와 n차원 속성 벡터 $\vec{x}$을 입력받아 $\vec{x}$에 대한 암호문(CT)과 인증키(K)를 생성하는 알고리즘이다.
- **디캡슐화(Decap)** $(TK_{\vec{\sigma}}, CT)$: 토큰 $TK_{\vec{\sigma}}$와 암호문 CT를 입력으로 받아 $f_{\vec{\sigma}}(\vec{x})=1$을 만족하는 경우 인증키 K를, $f_{\vec{\sigma}}(\vec{x})\neq 1$인 경우 $\perp$를 출력하는 알고리즘이다.

정확성(correctness). 임의의 속성 벡터 $\vec{\sigma}$와 $\vec{x}$에 대하여, $(PK, SK) \xleftarrow{R} \mathbf{Setup}(k,n)$, $TK_{\vec{\sigma}} \xleftarrow{R} \mathbf{GenToken}(SK, \vec{\sigma})$, $(CT, K) \xleftarrow{R} \mathbf{Encap}(PK, \vec{x})$ 이라 하자. 만약 $f_{\vec{\sigma}}(\vec{x})=1$을 만족한다면, $\mathbf{Decap}(TK_{\vec{\sigma}}, CT)$는 K를 출력한다. 만약 $f_{\vec{\sigma}}(\vec{x})\neq 1$인 경우, 의미 없는(negligible) 값 $\epsilon(k)$에 대하여 $\Pr[\perp \leftarrow \mathbf{Decap}(TK_{\vec{\sigma}}, CT)] > 1-\epsilon(k)$ 이다.

동일성 술어(Equality predicate). 제안하는 인증 프로토콜 설계에 사용되는 HV-KEM은 토큰과 암호문에 사용되는 두 개의 속성 벡터 $\vec{\sigma}, \vec{x}$에 대하여 아래에서 정의되는 매칭 조건 함수 $f_{\vec{\sigma}}$로 동일성 술어를 표현한다. 함수 $f_{\vec{\sigma}}$는 다음과 같이 정의된다.

$$f_{\vec{\sigma}}(\vec{x}) = \begin{cases} 1, \text{ if for } all\ i \in [1,n],\ \sigma_i = x_i, \\ 0, \ otherwise. \end{cases}$$

2.2 보안 요구사항

본 절에서는 사용자 인증 프로토콜 설계에 사용되는 HV-KEM 기법의 보안 요구사항을 정의한다. 두 가지로 (1) 속성 정보 프라이버시와 (2) 암호문 위조불가능성을 고려한다.

HV-KEM 기법이 보안 요구사항을 정의하기 위해 챌린저와 공격자가 서로 정보를 주고받는 가상의 게임(simulation) 환경을 이용한다. 챌린저 S는 공격자를 이용하여 자신에게 주어진 암호학적으로 어려운 문제(cryptographic hardness problem)를 풀 수 있어야 하고, 가상의 게임 환경에서는 공격자가 실제 공격 환경과 가상의 게임을 구분하지 못하도록 증명을 구성한다.

2.2.1 속성 벡터 프라이버시

속성 벡터 프라이버시는 공격자[1]가 공격 대상이 되는 속성 벡터로 만들어진 토큰과 암호문을 받더라

도 해당 속성 벡터에 대한 정보를 알아내는 것이 어렵다는 것을 의미한다. 여기서 공격자는 사용자의 공개키뿐만 아니라 속성 벡터로 만들어진 토큰을 받는다. 또한 속성 벡터로 생성된 정당한 암호문 집합을 얻을 수 있다. 이러한 정보에도 불구하고 속성 벡터 프라이버시가 보장되려면 공격자는 사용자의 속성 벡터에 대한 정보를 아는 것이 어려워야 한다. 제3장의 사용자 인증 프로토콜에서 속성 벡터는 생체 정보(biometric data)로 대체될 것이다.

공격자에게 속성 벡터의 차원 n이 주어진다고 하자. 공격자 A는 챌린저 S와 다음과 같이 가상의 게임을 수행한다.

- ***Init.*** S는 공격 대상이 되는 속성 벡터 x^*를 선택하고, 공개키 PK와 비밀키 SK를 생성한다. 그리고 비밀키 SK와 속성 벡터 x^*를 이용하여 토큰 TK_{x^*}를 생성한다. S는 A에게 공개키 PK와 토큰 TK_{x^*}를 전송한다.
- ***Query.*** A는 S에게 암호문을 질의하고, S는 속성 벡터 x^*와 공개키 PK를 이용하여 생성된 암호문 CT를 전송한다. A의 암호문 질의 개수는 다항식 시간 내에 계산할 수 있는 것으로 제한된다.
- ***Output.*** A는 ***Init***에서 S가 선택한 공격 대상이 되는 속성 벡터를 추측하여 σ^*를 S에게 전송한다.

공격자 A가 속성 벡터 x^*와 동일한 값을 추측하면 A가 게임을 이긴 것(win)으로 간주한다. 즉, A가 게임을 이긴다는 것은 $x^* = \sigma^*$이 성립함을 의미한다. 이 때, 게임(PA: Privacy of Attribute data)에서 공격자 A가 얻는 이점(advantage)은 다음과 같이 정의한다.

$$Adv_A^{PA} = \Pr[A\ wins]$$

정의 2. HV-KEM 기법에 대한 임의의 다항식 시간 공격자 A에 대하여 공격자의 이점 Adv_A^{PA}이 의미 없는(negligible) 값이라면, HV-KEM 기법은 속성 벡터 프라이버시 공격에 대해 안전하다.

1) 여기서의 공격자는 제3장의 HV-KEM을 이용한 인증 프로토콜에서 인증서버를 말한다.

속성 벡터 프라이버시 게임에서 생기는 자연스러운 의문은 공격자가 사용자의 공개키와 토큰을 받으면서도, 토큰에 포함된 속성 벡터를 어떻게 숨길 수 있는가 하는 것이다. 당연히 공격자는 자신이 생각한 속성 벡터를 선택하여 (공개키가 있으므로) 암호문을 생성할 수 있고, 이를 토큰으로 디캡슐화하여 정상적인 인증키가 나오면 토큰에 포함된 속성 정보 (또는 일부)를 역으로 추측할 수 있기 때문이다. 이러한 공격은 속성 벡터의 차원 n이 낮다면(즉, n이 작다면) 다항식 시간 내에 성공하게 된다. 따라서 다항식 시간 내의 공격에 견디려면 기본적으로 n이 커야 한다. 예를 들어 n이 100이라 하고, 각 속성 벡터의 성분이 0과 1로 표현된다면, 성공적인 추측을 위해서는 최악의 경우로 2^{100}의 암호문 작성 및 디캡슐화 연산을 수행해야 한다. 이하에서는 속성 벡터의 프라이버시를 언급할 때 속성 벡터의 차원 n이 다항식 시간 내의 추측 공격을 견디는 값으로 설정되었다고 가정한다.

2.2.2 암호문 위조불가능성

암호문 위조불가능성은 공격자가 사용자의 속성 정보 없이도 정상적으로 디캡슐화가 되는 인증키를 생성할 수 없어야 한다는 것을 의미한다. 여기서 공격자[2]는 사용자의 공개키를 받고, 속성 정보로 생성된 토큰은 받지 않는다. 일단 공개키가 주어지므로 공격자가 선택한 임의의 속성 벡터를 이용하여 원하는 암호문과 키를 생성할 수 있다. 또한 공격자에게는 공격 대상 사용자의 속성 벡터로 생성된 정당한 암호문 집합이 주어진다. 이러한 환경에서 공격자의 목표는 주어진 정보를 이용하여 공격 대상 사용자가 정상적으로 생성했던 (이전의 암호문 중 하나에 대응하는) 인증키를 생성하는 것이다. 이를 가상의 게임으로 정의하면 다음과 같다.

공격자 A에게 속성 벡터의 차원 n이 주어진다고 하자. 공격자 A는 챌린저 S와 다음과 같이 암호문 위조불가능성 게임을 수행한다.

- ***Init.*** S는 공격 대상이 되는 속성 벡터 x^*를 선택하고, 공개키 PK와 비밀키 SK를 생성한다. S는 A에게 공개키 PK를 전송한다.

2) 여기에서의 공격자는 제3장의 HV-KEM을 이용한 인증 프로토콜에서 인증서버가 아닌 제3자를 말한다.

- ***Query.*** A는 S에게 암호문을 질의하고, S는 속성 벡터 x^*와 공개키 PK를 이용하여 생성된 암호문 CT를 전송한다. A의 암호문 질의 개수는 다항식 시간 내에 계산할 수 있는 것으로 제한된다.
- ***Output.*** A는 ***Query***에서 받은 기존 암호문 중 하나인 CT와 그에 해당하는 인증키 K를 생성하거나, 새로운 암호문 CT^*과 대응하는 인증키 K^*를 생성하여 S에게 전송한다.

공격자 A가 공격 대상 속성 벡터 x^*에 대응하는 기존 암호문 중 하나인 CT와 그에 대응하는 키 인증키 K를 생성하거나, 새로운 암호문 CT^*과 대응하는 인증키 K^*를 생성하면 A가 위 게임을 이긴 것(win)으로 간주한다. 즉, A가 게임을 이긴다는 것은, 비밀키 SK와 공격 대상 속성 벡터 x^*를 이용하여 생성된 토큰 TK_{x^*}에 대하여 $\mathbf{Decap}(TK_{x^*}, CT) = K$ 또는 $\mathbf{Decap}(TK_{x^*}, CT^*) = K^*$이 성립함을 의미한다. 이 게임(CU: Ciphertext Unforgeability)에서 공격자 A가 얻는 이점(advantage)을 다음과 같이 정의한다.

$$Adv_A^{CU} = \Pr[A \ wins]$$

정의 1. HV-KEM 기법에 대한 임의의 다항식 시간 공격자 A에 대하여 공격자의 이점 Adv_A^{CU}이 의미 없는(negligible) 값이라면, HV-KEM 기법은 암호문 위조불가능성 공격에 대해 안전하다.

III. 제안하는 생체 정보 기반 인증 프로토콜

본 장에서는 HV-KEM(Hidden Vector Key Encapsulation Mechanism)을 이용한 생체 정보 기반의 사용자 인증 프로토콜을 제안한다. 제안하는 프로토콜을 위해 다음과 같은 가정을 한다. (1) 사용자의 생체 정보(biometric data)는 n차원의 속성 벡터로 표현된다. (2) 속성 벡터의 차원 n은 다항식 시간 내의 추측 공격에 대항할 수 있는 값이라고 하자. (3) 사용자의 생체 정보를 추출하는 인지(recognition) 알고리즘이 있다고 가정하자. 단, 인지 알고리즘에는 암호학적으로 필요한 비밀키 관리 등의 조건을 요구하지 않는다. (4) 생체 정보는 인지 알고리즘에 의해 추출할 때마다 일정 확률 이하로 원래의 측정값과 다를 수 있다고 가정한다.

제안하는 생체 정보 기반 인증 프로토콜에서 핵심 사항은 (a) 사용자로부터 추출된 생체 정보가 인증서버 및 제3자에게 노출되지 않는다는 것과 (b) 매번 생체 정보 측정 시 (노이즈 등의 에러(error)로 인해) 일정 확률 이하로 원래 값과 다르게 변할 수 있음에도 사용자 인증이 정상적으로 이루어져야 한다는 것이다. 첫 번째 핵심사항은 HV-KEM의 안전성으로부터 쉽게 얻어질 수 있으나, 두 번째 핵심사항은 동일성 술어(equality predicate)를 지원하는 HV-KEM으로부터 해결하는 것이 쉽지 않아 보인다. 이를 해결하는 아이디어는, 간단히 설명하면, 사용자의 생체 정보 집합으로부터 다수의 생체 정보 부분집합을 구한 뒤, 각 부분집합에 대응하는 토큰을 생성하고, 인증 시 부분집합에 대응하는 토큰으로 검증이 통과되면 사용자를 인증하는 것으로 판단하는 것이다. 이는 *일부 부분집합 이외의 원소에서 원래의 생체 정보와 다른 값이 측정되더라도, 그 다른 값과는 관계없이 인증을 통과시킬 수 있는 효과가 있다.*

3.1 용어 정리(Notation)

제안하는 사용자 인증 프로토콜에서 생체 정보(속성 벡터) $\overrightarrow{w}$는 n차원이며, $\overrightarrow{w} = (w_1, \ldots, w_n)$으로 표현한다. 이 때, $\overrightarrow{w}$과 크기가 $l(<n)$인 부분집합 $R_i \subset \{1, \ldots, n\}$에 대하여 $\overrightarrow{w_{R_i}}$를 다음과 같이 정의한다.

$$\overrightarrow{w_{R_i}} = \{w_j \mid j \in R_i\}$$

제 2.1장에서 정의한 HV-KEM의 알고리즘을 각각 *HV-KEM.Setup, HV-KEM.GenToken, HV-KEM.Encap, HV-KEM.Decap* 으로 표현한다.

본 인증 프로토콜은 사용자와 인증서버 사이에서 이루어지며, 사용자가 인증서버에게 사용자임을 증명하는 과정을 다룬다. 기본적으로 인증 프로토콜은 도전-응답(challenge-response) 방법을 이용한다. 즉 인증서버가 보내온 메시지에 대해 사용자가 그에 대응하는 응답을 보내서 사용자를 인증한다. 인증 과정에서는 메시지 인증 코드(Message Authentication Code, MAC)를 사용하는데, MAC은 다음과 같이 두 개의 알고리즘으로 구성된다.

- **생성(Gen)**(K, m) : 대칭키 K와 메시지 m을 입력으로 받아 메시지 인증 코드(MAC)를 생성

하는 알고리즘이다.

- **검증(Verify)** (K, m, MAC^*) : 대칭키 K와 메시지 m, 메시지 인증 코드 MAC을 입력으로 받아 검증한다. $MAC^* \leftarrow \mathbf{Gen}(K, m)$을 만족하는 경우 1을, 아닌 경우 0을 출력한다.

MAC을 구성하는 두 개의 알고리즘을 각각 $MAC.Gen$, $MAC.Verify$ 로 표현한다.

3.2 제안하는 인증 프로토콜

제안하는 인증 프로토콜은 등록 단계와 인증 단계로 이루어진다.

3.2.1 등록 단계(Registration phase)

등록 단계에서는 사용자가 자신의 생체 정보를 이용하여 공개키와 토큰을 생성하고, 이를 인증 서버에 등록한다.

- *Step 1.* : 사용자는 보안 상수 k와 생체 정보 차원 n을 입력으로 받아 다음과 같이 공개키 PK와 비밀키 SK를 생성한다.

$$(PK, SK) \leftarrow \mathbf{HV\text{-}KEM.Setup}(k, n)$$

- *Step 2.* : 사용자는 인지 알고리즘을 이용하여 자신의 생체 정보 $\overrightarrow{w}$를 측정하고, $\overrightarrow{w}$로부터 다음과 같이 d개의 부분집합에 대응하는 토큰을 생성한다.
  (1) 랜덤하게 크기가 l[3]인 $\{1, \ldots, n\}$의 부분집합 R_i를 선택한다.
  (2) $\overrightarrow{w_{R_i}} = \{w_j \mid j \in R_i\}$에 대하여 다음과 같이 토큰 $TK_{\overrightarrow{w_{R_i}}}$를 생성한다.

$$TK_{\overrightarrow{w_{R_i}}} \leftarrow \mathbf{HV\text{-}KEM.GenToken}\left(SK, \overrightarrow{w_{R_i}}\right)$$

사용자는 다음과 같이 생체 정보 $\overrightarrow{w}$에 대응하는 토큰 $TK_{\overrightarrow{w}}$를 설정한다.

$$TK_{\overrightarrow{w}} = \left(TK_{\overrightarrow{w_{R_1}}}, \ldots, TK_{\overrightarrow{w_{R_d}}}\right)$$

- *Step 3.* : 사용자는 자신의 ID와 공개키 PK, (R_i를 포함한) 토큰 $TK_{\overrightarrow{w}}$를 인증 서버에 전송한다. 전송 후 사용자는 자신이 생성한 모든 값들을 삭제한다.

- *Step 4.* : 인증 서버는 사용자로부터 전송 받은 $\left(ID, PK, TK_{\overrightarrow{w}}\right)$을 DB에 저장한다.

3.2.2 인증 단계(Authentication phase)

인증 단계에서는 사용자와 인증서버가 HV-KEM을 이용하여 도전-응답 방식으로 인증을 수행한다. 인증을 수행하기 위해 사용자는 인지 알고리즘을 이용하여 자신의 생체 정보를 측정한다. 여기서 측정된 생체 정보는 등록 단계에서 측정된 생체 정보와 다를 수 있으며, 인증이 통과되기 위해서는 프로토콜 설계에서 허용하는 오차 범위 안에 있으면 인증을 통과하게 된다.

- *Step 1.* : 사용자는 인증서버에 인증 요청 메시지와 자신의 ID를 전송한다.

- *Step 2.* : 인증서버는 DB에서 ID에 대한 공개키 PK와 도전 값으로써 랜덤하게 생성된 메시지 m을 사용자에게 전송한다.

- *Step 3.* : 사용자는 전송받은 공개키 PK와 인지 알고리즘으로 새롭게 측정한 자신의 생체 정보 $\overrightarrow{w'} = (w_1', \ldots, w_n')$ 을 입력으로 받아 다음과 같이 응답 값을 생성한다.
  (1) 암호문 CT와 인증키 K를 생성한다.

$$(CT, K) \leftarrow \mathbf{HV\text{-}KEM.Encap}\left(PK, \overrightarrow{w'}\right)$$

  (2) 인증키 K를 이용하여 도전 값 m과 자신이 생성한 암호문 CT에 대한 메시지 인증 코드 MAC을 생성한다.

$$MAC \leftarrow \mathbf{MAC.Gen}(K, m \| CT)$$

- *Step 4.* : 사용자는 (CT, MAC)을 응답 값으로 하여 인증 서버에게 전송한다.

- *Step 5.* : 인증서버는 ID에 대응하는 토큰 $TK_{\overrightarrow{w}}$를 이용하여 다음과 같이 사용자로부터 전송받은 응답 값을 검증한다.

3) 여기서 부분집합의 크기 l은 제2장에서 언급한 속성 벡터의 추측 공격에 저항성을 가지도록 설정되어야 한다.

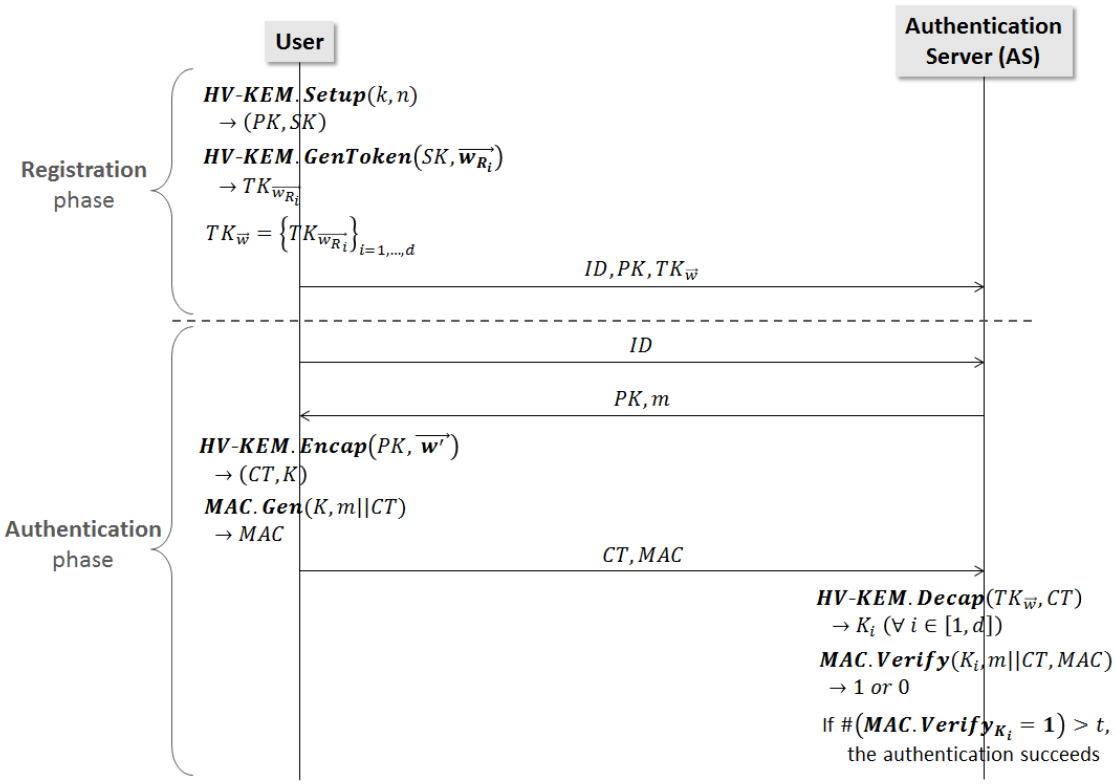

Fig. 1. Biometric authentication protocol using HV-KEM

(1) 모든 $i \in [1,d]$에 대해 키 K_i를 생성한다.
$K_i \leftarrow \text{HV-KEM.Decap}\left(TK_{\overrightarrow{w_{R_i}}}, CT\right)$

(2) 인증키 $K_i (i \in [1,d])$를 이용하여 전송받은 메시지 인증 코드 MAC을 검증한다.
$1 \text{ or } 0 \leftarrow \text{MAC.Verify}(K_i, m \| CT, MAC)$

(3) d개의 토큰 중 임계치(threshold) t개 이상이 검증에 통과하면 인증 서버는 사용자가 인증에 성공한 것으로 판단한다.

IV. 제안하는 인증 프로토콜 분석

4.1 인증 프로토콜의 정확성 분석

제안하는 인증 프로토콜은 등록 단계에서 n차원 벡터로 측정되는 생체 정보 $\overrightarrow{w}$와 인증 단계에서 n차원 벡터로 측정되는 생체 정보 $\overrightarrow{w'}$ 사이에 일정 확률 p이하로 오차가 발생하더라도 사용자 인증이 정상적으로 통과되어야 한다. 이러한 에러 처리는 n개 중 l개의 원소로 이루어진 부분집합을 d개 선택하고, 각 부분집합에 대응하는 생체 정보를 이용하여 d개의 토큰을 생성한 후, 일정한 임계치 t이상의 토큰에서 MAC 검증이 되는 정상적인 인증키가 생성되면 된다. 여기서 부분집합의 크기 l은 디캡슐화를 이용한 생체 정보의 추측 공격에 저항성을 가지도록 설정되어야 함을 상기하자.

구체적으로 오차확률이 최대 p라고 할 때, 설정되어야 할 d와 t값은 (n,l,d,t)의 변수들 간의 관계로 결정된다. 오차확률이 최대 p이면, 전체 n개의 원소 중 최대 pn개의 원소들이 다르게 측정될 수 있다. 이 경우 l개의 부분집합의 원소로 구성된 하나의 토큰이 인증을 통과하려면, 최대 pn개의 (에러가 발생한 각각의) 원소들이 전체 n개의 원소 중 이미 선택된 l개의 부분집합 원소가 아닌 나머지 $(n-l)$개의 원소들 중 하나이어야 한다. 이것은 *적어도* 아래의 확률로 하나의 토큰이 인증을 통과할 수 있음을 의미한다.

$$\frac{(n-l)}{n}\times\frac{(n-l-1)}{(n-1)}\times\cdots\times\frac{(n-l-pn+1)}{(n-pn+1)} \quad (1)$$

식 (1)에서 확률이 결정되면, 자연스럽게 토큰의 개수 d와 임계치 t가 결정된다. 예를 들어 하나의 토큰이 통과될 확률이 0.013이라고 하면, $d=1000$이고 $t=13$으로 결정된다. 이것은 최소 1000개의 토큰 중에 13개 이상 정상적으로 통과되면 사용자 인증으로 간주한다는 의미이다. 〈Table 1〉에서는 구체적인 수치로 토큰의 개수 d와 임계치 t를 계산한 사례이다.

첫 번째 행은 $n=200$, $l=80$, $p=0.05$이고, 하나의 생체 정보 원소(w_i)는 0과 1로 구성된 1bit 정보를 표현하는 경우이다. 80개의 원소로 구성된 생체 정보를 디캡슐화로 추측하기 위해서는 총 (최악의 경우) 2^{80}의 시도를 해야 한다. 이것은 다항식 시간 내에 충분이 안전한 생체 정보 구성이라고 할 수 있다. 이 경우 하나의 토큰이 인증을 통과하기 위한 최소 확률을 식(1)로 결정하면, 약 0.005가 된다. 따라서 $d=1000$, $t=5$로 설정하면, 등록 단계의 생체 정보와 인증 단계의 생체 정보가 최대 오차확률 $p=0.05$이하에서 다르게 되는 경우 사용자 인증이 통과된다. 두 번째 행은 최대 오차 확률 p가 커질 때, 훨씬 더 많은 토큰의 개수 d가 필요함을 보인다. 세 번째와 네 번째 행은 하나의 생체 정보 원소(w_i)가 {00,01,10,11}로 구성된 2bits 정보를 표현하는 경우이다. 동일한 오차 확률 p에 대해서 벡터의 크기 n이 커질수록 디캡슐화에 필요한 연산이 줄어드는 것을 볼 수 있다. 실제로 지문, 홍채 등의 생체 정보를 비트 단위의 스트링으로 추출하여 인증에 사용할 수 있으며, [13, 14]의 결과를 따르면 $(n,p,l)=(2048,\ 0.1,\ 80)$와 같이 표현할 수 있다.

Table 1. Numeric examples of (n,p,l,d,t)

$\|w_i\|$	n,p,l	d,t
1 bit	$n=200,\ p=0.05,$ $l=80$	$d=1000,$ $t=5$
1 bit	$n=200,\ p=0.1,$ $l=80$	$d=1000000,$ $t=18$
2 bits	$n=100,\ p=0.05,$ $l=40$	$d=100,$ $t=7$
2 bits	$n=200,\ p=0.05,$ $l=40$	$d=10,\ t=1$

구체적인 확률 p와 (n,l,d,t)의 값은 인증 프로토콜에 사용되는 지문, 홍채 등 생체 정보에 따라 결정될 것이다.

4.2 인증 프로토콜의 안전성 분석

제안하는 인증 프로토콜은 2.2장에서 정의한 속성 벡터 프라이버시와 암호문 위조불가능성을 만족하는 HV-KEM을 이용하여 설계한다. 이를 통해 다음의 공격자 유형으로부터 인증 프로토콜의 안전성을 보장할 수 있다.

4.2.1 악의적인 인증서버

인증서버는 등록 단계에서 사용자의 공개키 및 생체 정보($\overrightarrow{w}$)로부터 생성된 토큰($TK_{\overrightarrow{w}}$)을 얻을 수 있다. 또한, 인증 단계에서 사용자의 생체 정보($\overrightarrow{w'}$)로부터 생성된 암호문(CT)과 그에 대응하는 키를 쉽게 얻을 수 있다. 따라서 인증서버가 특정 사용자인 척하는 위장공격(impersonation attack)은 매우 쉽게 이루어진다. 제안하는 프로토콜에서는 인증서버가 위장공격을 하지 않는다고 가정하자. 이 경우 악의적인 인증서버의 공격목표는 주어진 정보로부터 사용자의 생체 정보를 얻는 것이다. 즉 인증 프로토콜이 악의적인 인증서버에 대해 안전하려면, 사용자로부터 전송받은 공개키, 토큰과 암호문으로부터 사용자의 생체 정보를 얻는 것이 어려워야 한다.

제안하는 인증 프로토콜에서 인증서버가 얻는 값은 사용자에 대한 공개키, 토큰의 집합, 그리고 사용자의 생체 정보로부터 만들어진 암호문들이다. 토큰이 하나라면, 이것은 제 2.2.2절에서 정의한 속성 벡터의 프라이버시를 정의하는 게임에서 공격자가 취할 수 있는 값들이다. 그러므로 HV-KEM 기법이 속성 벡터의 프라이버시를 보장한다면, 악의적인 인증서버는 하나의 토큰 및 주어진 정보로부터 사용자의 생체 정보를 얻을 수 없다. 그리고 토큰의 집합에 대해서는 위에서 설명한 속성 벡터 프라이버시 게임을 연속적으로(즉 하이브리드 게임의 증명 테크닉을 적용하여) 적용함으로써 토큰의 집합이 주어진 경우에도 속성 벡터의 프라이버시 게임에서 안전성을 보장할 수 있다. 결국 인증 프로토콜에서 사용된 HV-KEM 기법이 속성 벡터의 프라이버시를 보장한다면, 악의적인 인증서버가 사용자의 생체 정보를 취득하려는 공격에 안전하다.

Table 2. Characteristic comparison between previous authentication methods and ours

	OTP	Certificate	ID/PW	Fuzzy Extractor	Ours
Necessary information	OTP	certificate, password	password	biometric data, helper data	biometric data
What you **have**	O	O	X	△	X
What you **know**	X	O	O	X	X
What you **are**	X	X	X	O	O

4.2.2 악의적인 외부 공격자

여기서 외부 공격자는 서버와 인증서버를 제외한 공격자를 말한다. 악의적인 외부 공격자의 공격 목표는 공격 대상 사용자를 가장하여 인증서버에게 공격 대상 사용자로 인증을 통과시키는 위장공격이다. 이러한 공격은 HV-KEM의 암호문 위조불가능성에 의해 무력화되는데, 그 이유는 다음과 같다. 먼저 외부 공격자는 공격 대상 사용자의 ID를 쉽게 얻을 수 있다고 가정하자.[4] 이 경우 외부 공격자는 인증서버에 공격 대상 사용자의 ID를 전송함으로써 ID에 대응되는 공개키 PK를 쉽게 얻는다. 그리고 인증 단계에서 공격 대상 사용자와 인증서버 간에 주고받는 값을 수집하여, 공격 대상 사용자의 생체 정보($\overrightarrow{w'}$)로부터 생성된 암호문(CT)들과 대응되는 MAC값들을 얻을 수 있다. 안전한 MAC을 가정하면, 외부 공격자가 얻어내는 값들은 공개키와 암호문들의 집합이 되는데, 이 값들은 암호문 위조불가능성을 정의하는 안전성 게임에서 공격자가 취할 수 있는 값들이 된다. 여기서 공격자가 주어진 값들을 이용하여 위장공격을 성공하려면, (1) 새로운 암호문(CT^*)과 그에 대응되는 인증키(K^*)를 구할 수 있거나, (2) 이전 암호문들 중 하나에 대응하는 인증키를 구하면 된다. 이 값들 중 하나를 구하면, 공격자는 인증서버가 보내오는 메시지에 대해 정상적으로 검증이 통과되는 암호문 및 MAC값을 쉽게 생성할 수 있기 때문에 인증을 통과할 수 있다. 따라서 HV-KEM이 암호문 위조불가능성을 가진다면, (a) (새로운) 정상적인 암호문(CT^*)과 그에 대응되는 인증키(K^*)를 구하는 것이 어렵게 되고, (b) 이전 암호문들 중 하나에 대응하는 인증키를 구하는 것이 어렵게 된다. 이 두 경우 모두 외부 공격자의 위장 공격은 무력화된다.

제안하는 인증 프로토콜은 제 2.2.1절에서 정의한 암호문 위조불가능성을 보장하는 HV-KEM 기법을 이용하여 설계되었기 때문에 악의적인 외부 공격자의 위장공격으로부터 안전하다.

4.3 인증 프로토콜의 효율성 분석

제안하는 인증 프로토콜의 가장 큰 장점은 사용자가 등록 단계 이후 인증에 필요한 정보를 기억하거나 물리적으로 무엇인가를 안전하게 저장할 필요가 없다는 것이다. 사용자는 인증 단계에서 측정된 자신의 생체 정보를 이용하여 매우 편리하게 인증서버에게 사용자임을 증명할 수 있다. 여기서 생체 정보는 사용자 측에서 별도의 저장매체, 예를 들어, trust zone 같은 안전한 영역에 저장할 필요가 없다. 그리고 생체 정보는 HV-KEM에서 생성하는 토큰과 암호문에 포함되지만, 안전한 HV-KEM을 가정할 때 외부 공격자나 인증서버가 생체 정보를 획득하는 것은 어렵다. 또한 등록 단계에서 추출된 생체 정보와 인증 단계에서 추출된 생체 정보 사이에 일정 확률 이하로 오차가 발생하더라도 사용자 인증이 통과된다. 즉 일정 확률로 제시되는 오차 이내에서 생체 정보의 추출에서 발생하는 에러를 처리할 수 있다. 〈Table 2〉는 제안된 인증 기법과 기존의 인증 기법들을 비교한 결과를 나타내고 있다. 기호 '△'는 생체 정보에 대응하는 helper data를 사용자가 저장해야 함을 의미한다.

전체적으로 사용자와 인증서버의 계산적인 효율성은 토큰의 개수 d와 벡터의 차원 n에 반비례하게 된다. 사용자 측면에서는, 등록 단계에서 생체 정보

4) 만일 외부 공격자가 공격 대상 사용자의 ID를 얻는 것이 어렵다면, 제안한 인증 프로토콜의 안전성은 더 좋아질 것이다.

$\overrightarrow{w}$를 추출한 후 d개의 토큰을 생성해야 하고, 인증 단계에서는 생체 정보 $\overrightarrow{w'}$를 추출한 후 하나의 암호문과 MAC을 생성한다. 인증서버 측면에서는, 등록단계에서 사용자가 보내온 $(ID, PK, TK_{\overrightarrow{w}})$을 DB에 저장하고, 인증 단계에서는 사용자 보내온 암호문과 MAC값에 대하여 d개의 디캡슐화 과정을 수행하고 도출된 d개의 인증키로 MAC을 검증한다. 그러므로 구체적인 HV-KEM의 적용 시, 사용자와 인증서버 측면의 효율성을 전체적으로 고려하는 것이 필요할 것이다.

V. 결 론

본 논문에서는 HV-KEM을 이용하여 생체 정보 기반 인증 프로토콜을 제안하였다. 제안된 인증 프로토콜은 생체 정보의 프라이버시를 보장하고, 일정 확률 이하로 생체 정보의 오차를 허용하면서도 사용자를 인증할 수 있는 기법이었다. 사용자 측면에서는 매우 편리하면서도 안전한 인증 방법이었다.

향후 연구는 인증 프로토콜의 핵심 프리미티브로 사용된 HV-KEM 기법을 실제로 설계하고, 그 안전성을 증명하는 것이다. 기존에 제안된 HV-KEM 중 새롭게 정립된 안전성 모델 하에서 증명이 되는 것이 있는지 검토하는 것도 가능하고, 새로운 HV-KEM을 설계하는 것도 필요할 것이다. 예상되는 난관은 기존 HV-KEM의 안전성이 토큰을 제공하지 않는 모델에서 증명이 되었던 반면, 본 논문에서 정립한 안전성은 암호문 및 *토큰까지* 공격자에게 제공되는 모델에서 증명을 해야 하는 것이다. 이것은 기존 증명 이론과는 다른 접근 방법이 필요할 것이다.

References

[1] M. Abdalla and D. Pointcheval, "Interactive Diffie-Hellman assumptions with applications to password-based authentication," Proc. of the Financial Cryptography, LNCS 3570, pp. 341-356, Mar. 2005.

[2] M.R. Thompson, A. Essiari and S. Mudumbai, "Certificate-based authorization policy in a PKI environment," ACM Transactions on Information and System Security, vol. 6, no. 4, pp. 566-588, Nov. 2003.

[3] T.-C. Yeh, H.-Y. Shen and J.-J. Hwang, "A secure one-time password authentication scheme using smart cards," IEICE Transactions on Communications, vol. E85-B, no. 11, pp. 2515-2518, Nov. 2002.

[4] G.M. Ezovski and S.E. Watkins, "The electronic passport and the future of government-issued RFID-based identification," IEEE International Conference on RFID, pp. 15-22, Mar. 2007.

[5] "Apple pay contactless secure payment and tokenisation," (http://www.apple.com/apple-pay/)

[6] FUJITSU, "PalmSecure," (http://www.fujitsu.com/us/solutions/business-technology/security/palmsecure/palmsecure/)

[7] Y. Dodis, L. Reyzin and A. Smith, "Fuzzy extractor: how to generate strong keys from biometrics and other noisy data," Advances in Cryptology, EUROCRYPT'04, LNCS 3027, pp. 523-540, May 2004.

[8] Y. Dodis, B. Kanukurthi, J. Katz, L. Reyzin and A. Smith, "Robust fuzzy extractors and authenticated key agreement from close secrets," Advances in Cryptology, CRYPTO'06, LNCS 4117, pp. 232-250, Aug. 2004.

[9] A. Arakala, J. Jeffers and K.J. Horadam, "Fuzzy extractors for minutiae-based fingerprint authentication," Advances in Biometrics, ICB'07, LNCS 4642, pp. 760-769, Aug. 2007.

[10] Q. Li, M. Guo and E.-C. Chang, "Fuzzy extractors for asymmetric biometric representations," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1-6, Jun. 2008.

[11] D. Boneh, A. Sahai and B. Waters, "Functional encryption: definitions and

challenges," Proceedings of the 8th Theory of Cryptography Conference, pp. 253-273, Mar. 2011.
〔12〕 J.H. Park, "Efficient hidden vector encryption for conjunctive queries on encrypted data," IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 10, pp. 1483-1497, Oct. 2011.
〔13〕 C. Lee and J. Kim, "Cancelable fingerprint templates using minutiae-based bit-strings," Journal of Network and Computer Applications, vol. 33, no. 3, pp. 236-246, May 2010.
〔14〕 C. Chen and R. Veldhuis, "Extracting biometric binary strings with minimal area under the FRR curve for the hamming distance classifier," Signal Processing, vol. 91, no. 4, pp. 906-918, Apr. 2011.

〈저자소개〉

서 민 혜 (Minhye Seo) 학생회원
2012년 2월: 고려대학교 수학과 졸업
2012년 3월~현재: 고려대학교 정보보호대학원 석박사 통합과정
〈관심분야〉 암호 프로토콜, 인증 및 키 교환, 안전한 다자간 연산

황 정 연 (Jung Yeon Hwang) 정회원
1999년 2월: 고려대학교 수학과 졸업
2003년 2월: 고려대학교 정보보호대학원 석사
2006년 8월: 고려대학교 정보보호대학원 박사
2009년 5월~현재: 한국전자통신연구원 선임연구원
〈관심분야〉 암호이론, 프라이버시 강화 암호 기법, 바이오인증

김 수 형 (Soo-hyung Kim) 정회원
1996년 2월: 연세대학교 컴퓨터과학과 졸업
1998년 2월: 연세대학교 컴퓨터과학과 석사
2000년 11월: 한국정보통신연구원
2000년 12월~현재: 한국전자통신연구원 인증기술연구실 실장
〈관심분야〉 정보보호, 모바일결제, 핀테크

박 종 환 (Jong Hwan Park) 정회원
1999년 2월: 고려대학교 수학과 졸업
2004년 2월: 고려대학교 정보보호대학원 석사
2008년 8월: 고려대학교 정보보호대학원 박사
2013년 9월~현재: 상명대학교 컴퓨터과학과 조교수
〈관심분야〉 함수 암호, 브로드캐스트 암호, 암호 프로토콜