정 보 보 호 학 회 지
제27권 제3호, 2017. 6

바이오메트릭스 프라이버시 및 보안 기법에 대한 연구

김 혜 진*, 이 경 희**, 양 대 헌***

요 약

바이오메트릭스는 사용자가 가지는 신체적, 행동적 고유한 특성이기 때문에 유출 사고나 프라이버시 침해 사고에 매우 취약하다. 이러한 취약점을 보호하기 위하여 바이오메트릭스 시스템이 생성하는 템플릿은 반드시 원래의 데이터로 복원이 불가능하고, 취소 가능한 특성을 지녀야 한다. 본 논문에서는 바이오메트릭스 템플릿을 보호하는 특성을 부여하는 기법들을 분석하고, 특성을 제시하였다.

Ⅰ. 서 론

패스워드는 사용자가 알고 있는 정보(What you know)로써, 가장 보편적으로 이용되어 온 지식기반 인증 요소이다. 그러나 패스워드가 단순한 형태의 정보인 경우 탈취당할 위험성이 높아지고, 보안성을 위하여 복잡도를 높일 경우 사용자가 잊어버릴 가능성이 높아져 유용성이 떨어진다. 패스워드의 이러한 특징을 보완, 대체할 수 있는 요소로 각광받고 있는 요소가 특성기반 인증 요소(What you are)인 바이오메트릭스(biometrics)이다. 바이오메트릭스는 사용자의 고유한 신체적 혹은 행동적 특징으로, 별도의 암기나 소지 없이 인증에 활용할 수 있다는 장점을 가지고 있다. 최근 모바일 기술의 발전과 보편화로 인해, 모바일 기기를 활용한 바이오 인증 서비스가 점점 확산되어 가고 있다.

그러나 바이오메트릭스에는 패스워드와 같은 기존 인증 요소들과 다른 특성이 존재하고, 그에 따른 공격에 대한 위험성이 존재한다. 또한, 바이오메트릭스는 사용자의 개인정보로 취급되기 때문에 개인정보의 프라이버시를 보호해 줄 수 있어야 한다. 이러한 이유로, 바이오메트릭스의 보안과 프라이버시 보호에 대한 연구가 꾸준히 계속되고 있다. 본 논문에서는 바이오메트릭스의 종류 및 특성과, 바이오메트릭스의 보안과 프라이버시 보호 기법들에 대한 다양한 연구들을 소개한다.

Ⅱ. 바이오메트릭스의 특성과 보안

2.1. 바이오메트릭스의 특성 및 종류

바이오메트릭스는 사용자의 고유한 특성을 이용한 특성기반 인증 요소로, 잊어버리거나 분실할 위험이 적다. 바이오메트릭스는 크게 신체적 특징을 이용한 바이오메트릭스와 행동적 특징을 이용한 바이오메트릭스로 나누어 볼 수 있다. 사용자로부터 추출한 특징들을 가공해 디지털화 하여 인증 가능한 요소 형태로 만든 것을 바이오메트릭스 템플릿(template)이라고 한다.

신체적 특징을 이용하는 바이오메트릭스는 얼굴, 지문, 홍채, 장문 등 사용자의 신체부위로부터 고유한 특징을 추출하고 이를 비교하여 사용자를 인증하는 데 사용된다. 신체 부위마다 가지는 고유성에 정도의 차이가 존재하는데, 홍채와 같이 고유성이 높은 부위에서 뽑아낸 특징은 시간이 지나도 해당 정보가 사용자를 구별하는 성능을 유지한다. 이외에도 귀, 정맥 등 다양한 신체부위를 비롯하여 심전도, 뇌전도와 같은 생체신호로부터 고유성이 높은 특징을 추출하려는 시도가 지속되고 있다[10].

행동적 특징을 이용하는 바이오메트릭스는 사용자 행동의 고유한 특징을 추출하여 템플릿을 생성한다. 사용자의 걸음걸이, 목소리, 필기습관 등으로부터 나타나는 사용자의 고유한 습관 특징들을 추출하여 사용자를

* 인하대학교 컴퓨터공학과 (sunnyq@isrl.kr)
** 수원대학교 전기공학과 (khlee@suwon.ac.kr)
*** 인하대학교 컴퓨터공학과 (nyang@inha.ac.kr)

인증할 수 있다.

2.2. 바이오메트릭스의 보안

바이오메트릭스는 몇 가지 보안 및 프라이버시 이슈를 가지고 있다.[1]

우선, 비밀번호나 토큰 등 요소들은 오직 사용자만 알고 있거나, 가지고 있기 때문에 비밀로 유지가 되지만, 바이오메트릭스는 일상적으로 노출되기 때문에 비밀로 유지되는 요소가 아니다. 얼굴 사진이 찍히거나 물건에 묻은 지문 등으로 다양하게 유출되는 바이오메트릭스 정보를 이용하여 공격자가 인증을 시도할 수 있다.

두 번째로, 바이오메트릭스는 유출 사고에 매우 취약하다. 비밀번호의 경우, 공격자에 의하여 비밀번호 정보가 유출됐을 때 사용자의 비밀번호 등을 갱신하여 공격자가 해당 계정에 접근하는 것을 막을 수 있다. 그러나 바이오메트릭스 특징은 취소하거나 갱신할 수 없기 때문에, 정보가 유출될 경우 공격자는 유출된 정보를 이용하여 지속적으로 해당 사용자를 사칭할 수 있다.

마지막으로, 여러 인증 시스템에서 동일한 바이오메트릭스를 사용할 경우, 한 시스템에서 유출 되었을 때 다른 시스템에서도 유출 당한 것과 다름없다. 또한, 바이오메트릭스가 기관들 사이에서 공유될 경우, 템플릿을 통해 사용자를 추적할 수 있어 프라이버시에 위협이 될 수 있다.

바이오메트릭스는 이와 같이 내포하고 있는 보안 취약점이 많기 때문에 단일 요소로써 사용하기 보다는 2개 이상의 바이오메트릭스 요소를 결합하거나 패스워드나 토큰 등을 같이 사용하는 다중 요소(multi-factor) 인증 요소로 사용되도록 권장된다. 또한 템플릿으로부터 복원이 불가능하며(non-invertible), 취소 가능(cancelable)한 특성을 갖도록 해야 한다.

Ⅲ. 생체 인식 보안 기법

3.1. 복원 불가능한 바이오메트릭스

바이오메트릭스는 그 정보가 노출될 경우 해당 사용자가 바로 드러나거나, 노출된 데이터를 기반으로 타 시스템에서도 사용자를 추적할 수 있다. 이러한 이유로 바이오메트릭스의 경우 반드시 프라이버시의 보호가 필요하고, 템플릿으로부터 해당 사용자를 유추하거나 추적할 수 없도록 설계해야 한다.

3.1.1. Random Projection

랜덤 프로젝션은 고차원의 벡터 혹은 행렬을 기하학적 부분 공간(subspace)을 사상하여 저차원의 벡터 혹은 행렬로 변환하는 방법이다. 고차원의 공간에서 보이는 벡터끼리의 거리 차이가 저차원 공간에서도 유지되어, 고차원에서 사용자들끼리 보이던 차이를 저차원의 데이터로도 구현해 유지하여 사용자를 구분할 수 있다.

랜덤 프로젝션를 바이오메트릭스 템플릿과 곱하면 본래에 템플릿으로 되돌아갈 수 없는 복원 불가능한 성질을 갖는다. 또한, 같은 템플릿에 새로운 랜덤 프로젝션을 바꿔주면 결과가 달라지기 때문에 취소 가능한 특성도 함께 갖게 된다. 랜덤 프로젝션은 그 자체만 단독으로 사용하기 보다는 다른 여러 기법 안에 포함되어 사용되고 있다.

3.1.2. Random Permutation

랜덤 치환 기법은 바이오메트릭스로부터 특징을 추출한 특징 벡터 혹은 행렬에 랜덤으로 특징 값들을 섞거나 외부 값을 더하는 것을 의미한다. 템플릿의 값들이 랜덤하게 섞이게 되면 템플릿으로부터 원래의 바이오메트릭스 정보를 추출하기 어려워지기 때문이다.

J. Zuo 등 [5]은 특징 추출을 거친 홍채 템플릿에 shift와 XOR 연산을 이용하여 랜덤하게 템플릿을 섞는 기법을 제시하였다. 그런데 특징 추출 시 사용한 Gabor 기법이 그대로 선형 치환 기법을 거치게 되면 템플릿으로부터 많은 정보를 잃게 되고, 결국 인식 성능에 영향을 주게 된다.

위 기법의 단점을 보완할 수 있도록, 나누어진 해당 템플릿 부분 안에서 치환을 실행하는 기법[2]이 제시되었다. 이 기법은 홍채 이미지를 여러 구역(섹터)로 나누어 부분적으로 인식을 진행한 후 인식 결과 값을 종합하여 사용자를 인증한다. 홍채의 특징 벡터를 겹치지 않도록 섹터로 나누어 인식 알고리즘을 적용하는데, 각 섹터들이 특정 사용자의 라벨을 반환하게 된다. 이 라벨들 중 가장 다수를 차지하는 라벨을 통해 사용자를 인증한다. 이 기법은 투표 방식이기 때문에 일부 섹터가 노이

즈로 인해 다른 사용자로 오인식이 되어도, 노이즈가 적은 다수의 섹터들로 인하여 안정적인 인증 결과를 보이게 된다. 이러한 섹터 정보들을 랜덤 치환 기법을 이용하여 사전(dictionary)에 저장해 놓고 인식을 진행하기 때문에, 사전 정보가 없이 사전이 유출되어도 사용자의 바이오메트릭스 정보를 알기 힘들다.

또 다른 기법 중 하나인 General Permutation Transformation (GPT)[3]은 바이오메트릭스 이미지를 사상하여 특징 벡터를 만드는 프로젝션 행렬로부터 사용자의 본래 바이오메트릭스 정보를 복원할 수 있는 취약점을 보완하기 위하여 GPT라는 기법을 제안하였다. GPT는 사용자의 패스워드로부터 생성한 치환 행렬 P와 그의 역행렬 P^{-1}를 프로젝션 행렬과 얼굴 이미지 행렬에 곱하여 본래의 정보를 알아보기 힘들도록 만든다. 특징 벡터 생성 시, 이미지와 프로젝션 행렬에 각각 곱해져있던 P, P^{-1}가 단위행렬로 변환되며 본래의 특징 벡터가 생성된다. 사용자의 정확한 패스워드 없이는 정상적으로 특징 벡터의 생성이 어렵고, 프로젝션 행렬의 복원 역시 어렵기 때문에 유출 사고에 매우 강하다. 행렬의 연산 특성을 이용하여 사용자의 패스워드를 교체도 가능하다는 점에서 취소 가능한 특성을 갖는다.

3.2. 취소 가능한 바이오메트릭스

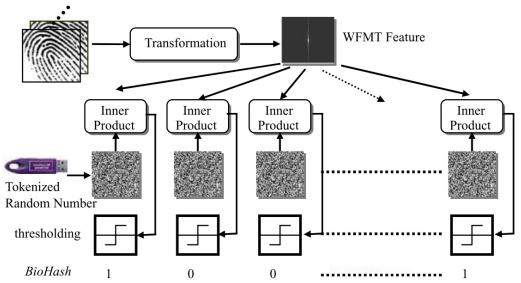
바이오메트릭스는 유출 사고 발생 시에도 변경이 불가능하기 때문에, 추출한 특징 그대로를 사용하지 않고 이를 가공하여 취소 가능한 템플릿의 형태로 사용해야 한다. 이 템플릿은 유출되어도 시스템에서는 기존 템플릿을 취소하고 새로운 템플릿을 생성할 수 있도록 설계되어야 한다.

3.2.1. 바이오해싱(BioHashing)

바이오해싱 기법[4]은 신체적 특징으로부터 안전한 비트스트링을 추출하는 이중 요소(two-factor) 기법이다. 추출한 비트스트링은 사용자 인증 혹은 바이오메트릭스 암호 시스템의 키로 사용이 가능하다.

바이오해싱은 얼굴과 지문으로부터 실수의 특징 벡터를 추출하고, 여기에 토큰을 이용한 랜덤한 숫자를 내적한 결과로부터 비트스트링을 생성한다. 비트스트링이 유출이 되어도, 이로부터 본래의 특징 벡터를 추출하기 어렵고, 토큰을 교체하면 새로운 비트스트링 생성이 가능하다.

지문을 이용한 BioHashing 기법에서는 Wavelet Fourier-Mellin 변환(WFMT)을 이용하여 복원 불가능한 특징 추출을 사용하였다, 이후 얼굴로부터 BioHashing을 생성하는 연구에서는 사용자별 토큰 대신 랜덤 프로젝션 기법을 사용하여 비트스트링을 생성 결과를 보였다. 그러나 각 사용자의 패스워드 혹은 토큰으로부터 랜덤한 숫자, 개별 랜덤 프로젝션을 생성한 후 바이오메트릭스 템플릿에 적용하는 경우에 패스워드나 토큰이 도난당하면 그 보안의 강도가 급격히 떨어지는 결과를 보이는 연구가 있었다[11].



〔그림 1〕 지문을 이용한 바이오해싱 기법의 기본 흐름〔4〕

3.2.2. Fuzzy 기반 기법

Fuzzy Commitment 기법〔6〕은 바이오메트릭스 템플릿을 이용하여 비밀 키를 숨기는 방법이다, 사용자가 등록한 템플릿과 동일하진 않더라도 유사한 템플릿을 입력하면, 이를 통하여 비밀 키를 복구할 수 있다. Fuzzy Commitment 함수에 비밀 키인 코드워드와 바이오메트릭스 템플릿을 이용해 코드워드를 숨긴다. 사전에 생성된 오프셋 정보를 이용하여 새로운 템플릿을 보정하여 등록 시 생성했던 템플릿과 유사하게 만들 수 있다.

Fuzzy Vault 기법〔7〕도 비밀 키를 숨기는 데 바이오메트릭스 정보를 이용한다는 점에서 Fuzzy Commitment 기법과 유사하다. 다항식 재구성 기법으로부터 착안한 이 기법은, 공간상에 비밀 키의 정보와 바이오메트릭스로부터 뽑아낸 정보를 섞어서, 정상 사용자의 바이오메트릭스가 들어오면 바이오메트릭스로 정보를 제거하여 비밀 키를 추출할 수 있게 한다. 다른 사용자의 바이오메트릭스가 들어올 경우, 해당

정보가 정상 사용자와 다르기 때문에 비밀 키에 해당하는 정보만 추출할 수 없어 인증이 불가능하다. Fuzzy Commitment와 Fuzzy Vault 기법과 같이 임의로 생성한 비밀 키 정보와 바이오메트릭스 정보를 혼합하는 기법을 key-binding이라 하고 바이오메트릭스를 이용한 암호화 시스템(biometric cryptosystem)을 설계할 때 적용한다.

반면, 바이오메트릭스 데이터를 키 생성의 시드로 생성하거나, 데이터 자체로부터 키를 추출하는 기법을 key-generation이라고 한다. Y. Dodis 등[12]은 노이즈가 있는 바이오메트릭스 정보를 키로 변환하여 암호화 시스템에 사용하거나 인증용으로 사용 가능한 Fuzzy Extractor 기법을 제시하였다. 바이오메트릭스를 이용해 랜덤한 비트스트링과 보조데이터를 생성하는데, 이후 등록 시 사용한 바이오메트릭스와 약간 다른 정보가 들어오더라도 보조데이터를 이용하여 동일한 랜덤 스트링을 생성할 수 있다.

3.2.3. 바이오토큰

바이오토큰[8]은 노이즈가 있는 바이오메트릭스 데이터를 두 부분으로 나누어 동일 사용자의 인증 시, 항상 안정적인 인증 결과가 나올 수 있도록 설계된 기법이다. 바이오메트릭스 데이터의 항상 안정적(stable)한 부분과 불안정(unstable)한 부분으로 분리하여 인증을 진행하게 된다. 안정적인 부분은 해시 혹은 PKI 방식으로 암호화하여, 인증 시 서버에 저장되어 있던 템플릿과 동일해야 하고, 불안정한 부분끼리는 거리 측정을 통하여 일정 범위 내이면 동일 사용자라고 판단한다. 안정적인 부분이 매칭이 안 될 경우, 불안정한 부분이 허용 범위 내에 있더라도 인증이 되지 않는다. 또한, 가변성이 존재하는 불안정한 부분과 일관적인 안정적 부분이 분리되어 있기 때문에 안정적인 부분을 여러 차례 암호화하여 복원 불가능하도록 템플릿을 보호할 수 있다. 이러한 기본 설계를 이용하여 양자 분리 바이오토큰(Bipartite biotokens) 인증 프로토콜을 함께 제시하였다.

3.2.4. 블룸필터(Bloom Filters)

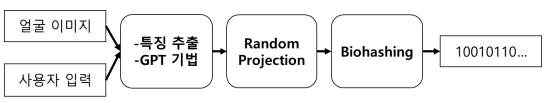
블룸필터는 공간 활용도가 높은 데이터 구조로, 원소와 집합간의 포함 관계를 나타낼 수 있다. C. Rathgeb 등[9]은 이 블룸필터를 이용하여 바이오메트릭스 보호, 템플릿 크기 압축 등의 효과를 가지는 인증 기법은 설계하였다. 우선, 홍채로부터 비트스트링 형식의 특징 벡터를 추출하고 이를 일정한 크기의 블록들로 나눈다. 각 블록 안의 비트들에 블룸필터를 적용하여 템플릿을 만들고, 모든 블록의 템플릿을 모아 최종 템플릿으로 사용한다. 블룸필터 기법으로 만들어진 템플릿은 회전된 이미지 등에 강하고, 사용자만 알고 있는 비밀 벡터를 템플릿에 XOR 시키는 방법을 통해 공격자가 본래의 템플릿을 추측하지 못하게 할 수 있다.

Ⅳ. Hybrid method

보안성 측면에서 바이오메트릭스를 단독 보안 요소로 사용할 경우에 보안성의 강도가 낮기 때문에, 패스워드나 토큰 등을 함께 사용하여 보안 강도를 높일 것이 권고된다. 기존의 제시되었던 기법들인 랜덤 프로젝션과 바이오해싱, GPT 기법을 융합하여 본래의 바이오메트릭스 템플릿의 복원이 어렵고, 취소가 가능한 바이오메트릭스 기반 이중 요소 키 생성 기법에 대하여 연구하였다.

얼굴 사진과, 패스워드 같은 사용자 입력정보를 이용하여 비트스트링 키를 생성할 수 있다. 얼굴로부터 PCA 혹은 LDA 기법을 이용하여 특징 정보를 추출한다. 그러나 특징 정보를 추출하기 위한 프로젝션 매트릭스는 사용자 입력으로부터 만들어진 랜덤 치환 행렬로 암호화되어 있기 때문에, 사용자 입력이 정확하지 않으면 올바른 특징 정보가 추출되지 않는다. 생성된 특징 정보를 RP 행렬을 적용한 후, 바이오해싱 기법을 이용하여 이진화를 거쳐 비트스트링 키를 생성한다. 바이오메트릭스의 노이즈는 ECC와 같은 보조 데이터를 생성하여 일부 보정이 가능하도록 하고, 비트가 일정 개수 이상 틀릴 때에는 동일 사용자가 아니라고 판단할 수 있다.

위 기법을 이용하여 비트스트링 키의 동일성을 판별하고 동일한 사용자임을 인증하는 인증 시스템으로 활



[그림 2] RP, 바이오해싱, GPT를 접목한 취소 가능한 바이오메트릭스 키 생성 기법

용하거나, 바이오메트릭스를 이용한 암호화 시스템에 적용가능하다. 사용자의 바이오메트릭스 템플릿은 저장되지 않아 유출 사고에 대한 위험이 적고, 랜덤 프로젝션 교환을 통해 키를 재발급할 수 있다.

Ⅴ. 결 론

바이오메트릭스는 이전부터 본인을 인증할 수 있는 확실한 보안 요소로 여겨져 왔으나 현실 세계에서 활용하기에는 많은 문제점이 존재했다. 이를 해결하기 위하여 데이터의 노이즈는 특징 추출 기법의 성능을 높여, 안정적으로 고유성이 높은 특징을 추출할 수 있는 기법들과, 바이오메트릭스 정보를 보호하도록 변형하고, 취소할 수 있는 기법들이 다양하게 제시되었다. 본 논문는 이러한 기법들에 특징과, 일부 기법을 복합한 새로운 기법에 대하여 연구하였다.

다양한 기법들이 제시되고 있지만, 이와 동시에 그에 대한 취약점들 역시 같이 연구되고 있다. 또한, 활용하는 바이오메트릭스의 특징에 따라 적용 방법 역시 각기 달라진다. 아직까지 사용자의 바이오메트릭스를 노출로부터 보호할 수 있는 궁극적인 방안은 없지만, 지속적인 연구 발전을 통하여 기존의 취약점을 보호할 수 있는 다양한 기법이 제시되고, 이를 통해서 실생활에 바이오메트릭스를 이용한 다양한 서비스들이 확산될 것으로 예상된다.

참 고 문 헌

[1] V.M. Patel, N.K. Ratha, and R. Chellappa, “Cancelable biometrics :A review,” *IEEE Signal Processing Magazine*, vol. 32(5), pp. 54-65, Sept. 2015

[2] J.K. Pillai, V.M. Ptel, R. Chellappa, and N.K. Ratha, “Secure and robust iris recognition using random projections and sparse representations,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33(9), pp. 1877-1893, 2011

[3] J. Kang, D. Nyang, and K.Lee, “Two-factor face authentication using matrix permutation transformation and a user password,” *Information Sciences*, 269, pp. 1-20, 2014

[4] A.T.B. Jin, D.N.C. Ling, and A.Goh, “Biohashing: two factor authentication featuring fingerprint data and tokenised random number,” *Pattern recognition*, vol. 37(11), pp. 2245-2255, 2014

[5] J. Zuo, N.K. Ratha and, J.H. Connell, “Cancelable iris biometric” *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on,* pp. 1-4, 2008

[6] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” *Proc. ACM Conf. Compter and Communications Security*, pp. 28-36, 1999

[7] A. Juels and M. Sudan, “A fuzzy vault scheme,” *Information Theory, 2002. Proc. IEEE International Symposium,* pp. 408, 2002

[8] W.J. Scheirer and T.E. Boult, “Bio-cryptographic protocols with bipartite biotokens,” *Biometrics Symposium*, BSYM’08, pp. 9-16, 2008

[9] C. Rathgeb F. Breitinger and C. Busch, “Alignment-free cancelable iris biometric templates based on adaptive bloom filters,” *Biometrics (ICB) 2013 International Conference,* pp. 1-8, 2013

[10] 김재성, 이새움, “생체신호를 이용한 텔레바이오 인식기술 동향 및 전망,” 정보보호학회지, 26(4), pp. 41-46, 2016

[11] A. Kong, K.H. Cheung, D. Zhang, M. Kamel, and J. You “An analysis of BioHashing and its variants,” *Pattern Recognition,* 39(7), pp. 1359-1368, 2006

[12] Y. Dodis, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” *Int. conference on the theory and applications of cryptographic techniques,* pp. 523-540, 2004

〈저 자 소 개〉

김 혜 진 (Hyejin Kim)
학생회원

2016년 2월 : 인하대학교 컴퓨터정보공학과 졸업
2016년 3월~현재 : 인하대학교 컴퓨터공학과 석사과정
관심분야 : 암호이론, 생체인증, 네트워크 보안

이 경 희 (KyungHee Lee)
정회원

1993년 2월 : 연세대학교 컴퓨터과학과 학사
1998년 8월 : 연세대학교 컴퓨터과학과 석사
2004년 2월 : 연세대학교 컴퓨터 과학과 박사
1993년 1월~1996년 5월 : LG 소프트(주) 연구원
2000년 12월~2005년 2월 : 한국전자통신연구원 선임연구원
2005년 3월~현재 : 수원대학교 전기공학과 부교수
관심분야 : 바이오인식, 정보보호, 컴퓨터비전, 인공지능, 패턴인식

양 대 헌 (DaeHun Nyang)
종신회원

1994년 2월 : 한구고가학기술원 과학기술대학 전기 및 전자공학과 졸업
1996년 2월 : 연세대학교 컴퓨터과학과 석사
2000년 8월 : 연세대학교 컴퓨터과학과 박사
2000년 9월~2003년 2월 : 한국전자통신연구원 정보보호연구본부 선임연구원
2003년 2월~현재 : 인하대학교 컴퓨터정보공학과 교수
관심분야 : 암호이론, 암호프로토콜, 인증프로토콜, 무선 인터넷 보안, 네트워크 보안