Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.28, NO.2, Apr. 2018 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2018.28.2.417

지능형 영상 감시 환경에서의 개인정보보호를 위한 COP-변환 기반 메타데이터 보안 기법 연구*

이 동 혁,[1†] 박 남 제[2‡]
[1]제주대학교 초등교육연구소, [2]제주대학교 초등컴퓨터교육전공

A Study on COP-Transformation Based Metadata Security Scheme for Privacy Protection in Intelligent Video Surveillance*

Donghyeok Lee,[1†] Namje Park[2‡]
[1]Elementary Education Research Institute, Jeju National University,
[2]Teachers College, Jeju National University

요 약

지능형 영상감시 환경은 CCTV 등에서 실시간으로 수집한 영상데이터의 분석을 통해 영상 객체에 대한 다양한 정보를 추출하고 이를 기반으로 자동화된 처리를 가능하게 하는 기술이다. 그러나 지능형 영상감시 환경에서는 프라이버시 노출 문제가 발생할 수 있어 이에 안전한 대책이 필수적이다. 특히, 영상 메타데이터에는 빅데이터 기반으로 분석된 다양한 개인정보가 포함될 수 있어 높은 위험성을 안고 있으나, 효율성의 문제로 영상메타에 암호화 방식을 적용하는 것은 적절하지 않다. 본 논문에서는 영상메타를 안전하게 보호할 수 있는 COP-변환 기법을 제안한다. 제안한 방식은 메타 변환을 통하여 원본 메타정보를 복원할 수 없도록 하며, 변환된 데이터에 직접적으로 질의를 가능하게 하므로 영상 메타데이터 처리과정에서의 안전성과 효율성을 크게 높인다는 장점이 있다.

ABSTRACT

The intelligent video surveillance environment is a system that extracts various information about a video object and enables automated processing through the analysis of video data collected in CCTV. However, since the privacy exposure problem may occur in the process of intelligent video surveillance, it is necessary to take a security measure. Especially, video metadata has high vulnerability because it can include various personal information analyzed based on big data. In this paper, we propose a COP-Transformation scheme to protect video metadata. The proposed scheme is advantageous in that it greatly enhances the security and efficiency in processing the video metadata.

Keywords: Video Surveillance, Privacy Protection, COP-Transformation, CCTV Video Security

I. 서 론

최근 지능형 영상감시 환경의 도입이 활발해지고 있다. 지능형 영상감시 기술은 CCTV 등에서 수집된 영상 정보를 수집 및 분석하여 이를 기반으로 자동화된 처리가 가능하게 하는 기술이다. 지능형 영상감시 기술은 사람, 자동차, 건물, 환경 등 다양한 분야에 적용할 수 있으며, 클라우드 및 빅데이터 분석

Received(09. 29. 2017), Modified(1st: 01. 29. 2018, 2nd: 02. 26. 2018), Accepted(03. 05. 2018)
* 이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임[2017-0-00207,클라우드 기반 지능형 영상보안 인큐베이팅 플랫폼 개발]
† 주저자, bonfard@jejunu.ac.kr
‡ 교신저자, namjepark@jejunu.ac.kr(Corresponding author)

기술과 결합하여 보다 의미 있는 상황 인식이 가능하다는 장점이 있다. 따라서 CCTV에서 입력되는 실시간 영상 분석을 통하여 보다 신뢰성 높은 위험요소 판단이 가능하며 이에 따른 적절한 조치를 취하는 것이 가능하게 됨으로써 지능형 영상감시 기술은 향후 지능치안 환경에서의 핵심 기술로서 작용하게 될 것으로 보인다[1].

지능형 감시 기술의 응용분야로써, 특정 건물이나 지역의 화재감지 발생의 예측 및 감시, 대기오염 등 환경 감시, 태풍/지진/해일 등 기후 감시, 자동차 교통안전 감시 등 다양한 적용분야가 있을 수 있으며, 이 중에서도 중요한 분야로 주목되는 부분은 행동인식 기반의 위험도 감지를 통한 범죄예방 분야가 있다. 현재 미국 뉴욕에서는 DAS(Domain Awareness System)이라는 지능형 범죄예방 시스템을 운영하고 있으며, 국내에서도 송도국제도시에서 IBM의 스마트 감시시스템인 SSS(Smart Surveillance Solution)을 운영하고 있다. 딥러닝 기술의 발달에 따라 이러한 지능형 영상감시 시스템은 CCTV 영상에 대한 방대한 빅데이터 기반의 축적된 데이터 분석을 기반으로 더욱 신뢰도 높은 행동인식, 위험요소 추론이 가능해질 것으로 보인다[2].

그러나, 이러한 기술 발달의 이면에는 다양한 위험성이 도사리고 있다. 예를 들어, 지능형 영상 감시 시스템이 해킹을 당하게 될 경우 심각한 사회적 문제가 될 수 있을 것이다. 여기에는 지능형 영상 감시 시스템 자체에 대한 오작동이나 마비를 일으키는 등 영상 감시 시스템 운영상의 다양한 위협이 발생할 수도 있으며, CCTV로부터 수집된 영상정보에 대한 노출이 발생할 수 있어 개인 프라이버시의 문제로 이어질 수 있다[3].

지능형 영상감시 시스템에서는 특정 개인의 위치, 이동 경로 등을 CCTV로부터 수집하여 클라우드 등 서버 스토리지 환경에 저장하게 될 것이며, 이러한 영상정보는 암호화하여 저장할 필요가 있다. 해킹 등에 의하여 영상정보가 노출되더라도, 공격자는 결국 암호화된 데이터만 습득할 수 있으므로 기밀성을 보장할 수 있다. 그러나, 영상정보를 암호화하여 저장하는 것은 간단한 해결책이 될 수는 없다. 영상정보는 대용량 데이터로써, 영상 데이터 분석시 다시 원본 영상으로 복호화하는 과정이 필요할 것이며, 여기에서의 오버헤드 문제가 매우 큰 걸림돌이 될 것이다. 이러한 문제를 해결하기 위해 모든 영상 데이터의 분석처리가 완료된 이후 해당 영상을 암호화를 해서 보관하는 방식의 정책으로 처리하더라도, 해당 영상이 분석된 메타데이터에 대한 보관 문제가 추가적으로 발생한다. 만약 영상 분석 메타데이터를 암호화하여 보관할 경우 동일한 복호화 오버헤드 문제가 발생할 것이며, 메타데이터를 평문으로 보관하게 될 경우는 해커가 메타데이터를 획득하는 것 만으로도 원본 CCTV 영상에 대한 상당한 정보를 얻을 수 있어 문제가 된다. 따라서, 본 논문에서는 CCTV 영상에 대한 개인정보보호를 위하여 COP(Character Order Preserving)-변환 기반의 메타데이터 보안 기법을 새롭게 제안한다. 본 논문에서 제안하는 COP-변환 방식은 평문 메타데이터의 정보를 전혀 다른 문자로 변환하여 원본 데이터를 식별할 수 없게 하는 방식이며, 평문의 문자열 순서 정보를 그대로 유지하고 있어 변환된 메타정보 자체만으로도 키워드 검색이 가능하다는 장점이 있다. 이러한 메타정보 COP-변환 방식을 기반으로, CCTV 영상데이터와 메타정보 모두 기밀성을 유지할 수 있어 개인 프라이버시 노출에 안전을 보장할 수 있다.

II. 관련 연구

본 장에서는 기존의 영상개인정보보호 기법을 살펴보고, 본 논문에서 위험요소로 지적하고 있는 영상 메타데이터 보안의 필요성에 대해 논의한다.

2.1 기존 영상개인정보보호 기법

2.1.1 영상 프라이버시 마스킹 기술

프라이버시 마스킹 기술은 일반적으로 영상의 얼굴 데이터를 알아볼 수 없도록 변경하는 것을 의미한다[4-6]. 예를 들어, 블러링, 픽셀화, 얼굴영상 제거 방식을 들 수 있는데, 이러한 방식은 근본적으로 복원이 필요한 경우에도 원본 영상으로 완전하게 복원할 수 없다는 한계점을 안고 있다.

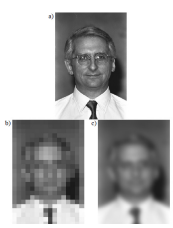


Fig. 1. Example of Privacy Masking Technology[6]

프라이버시 마스킹 기술은 구현의 용이성 및 프라이버시 보호의 장점이 있으므로 현재 다수의 CCTV 보안 제품에서는 마스킹 기법을 활용하고 있다. 그러나 향후 빅데이터 기반의 영상 복원기술인 Deep-Resolution 기술이 도입되면, 공개된 영상만으로도 블러링 및 픽셀화된 얼굴 정보가 원본에 가깝게 복원될 수 있을 가능성이 있어, 프라이버시 위험에서 안전을 보장할 수 없다[7,8].

일방향 프라이버시 마스킹 기법의 단점을 보완하기 위해 ROI(Region of Interest) 영역의 부분 암호화 방식도 제안된 바 있다[9]. ROI 부분암호화 방식은 영상 내의 얼굴 등 특정 영역만을 암호화하는 것으로, 통상적인 암호화 알고리즘과 동일하게 암호화 시는 얼굴정보를 식별할 수 없으나, 암호화 키를 활용하여 영상을 복호화할 경우 원본 영상을 복원할 수 있는 방식이다. 그러나 부분 암호화 방식은 암호화된 ROI 영역에 대한 메타정보를 어떻게 생성하고 보호할 것인지에 대해서는 언급하지 않고 있다. 구체적으로, 암호화된 영상 스토리지 상에서 특정인만 검색하여 원본 영상으로 복원하고자 할 경우는 실질적으로 처리할 수 있는 방법이 없으며, 이를 위해 임의로 검색 메타정보를 생성할 경우는 해당 메타정보에 그대로 영상 주체가 노출된다는 문제점이 존재한다.

2.1.2 클라우드 암호화 기반의 영상보안

D. A. Rodríguez-Silva 등은 클라우드 기반의 영상감시 환경을 제안하였다[10]. 지능형 영상 감시 환경은 대용량 데이터를 취급한다는 특성이 있어 확장성과 가용성이 필수적으로 요구되어 Amazon S3 기반의 확장 가능한 아키텍처를 제안하고 있다. 해당 아키텍처는 프라이버시 보호를 위해 SSL 프로토콜로 종단간 암호화할 것을 명시하고 있으며, 암호화 문제는 Amazon S3에서는 자체 암호화가 지원되므로 보안 문제를 해결하였음을 언급하고 있으나, 영상정보 처리와 저장방식에 대한 부분으로 한정되어 있다는데 한계점이 있다. 즉, CCTV 영상 데이터는 클라우드 환경에서 암호화되어 처리되지만, 해당 영상에 대해 어떤식으로 메타정보를 구성하고, 보호할 것인지, 암호화된 데이터에 대한 검색을 어떻게 처리할 것인지에 대해서는 언급하고 있지 않다. 실질적으로 이러한 방식은 향후 빅데이터 기반의 지능형 영상 감시 환경에 적용하기에는 한계가 있다[9-16].

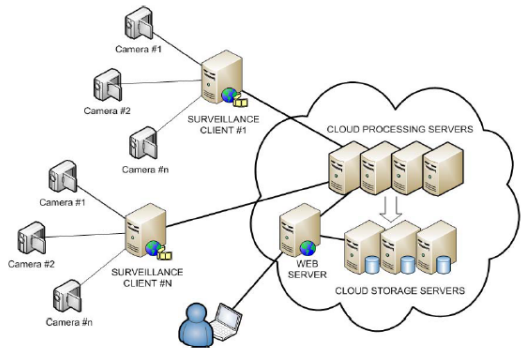


Fig. 2. Video surveillance based on cloud storage[10]

2.2 영상 메타데이터 보호의 필요성

2.2.1 영상 메타데이터에서의 개인정보 노출

향후 영상 감시 환경에서는 빅데이터를 기반으로 CCTV 영상데이터를 분석하여 특정 개인을 인식할 뿐만 아니라, 해당 객체의 현재 행동을 인식하며 행동 패턴을 분석하고 추론할 수 있게 될 것이다. 따라서 영상 메타데이터의 범위는 단순히 특정 개인의 신원을 인식하는 수준이 아니라, 개인에 대한 감정, 현재 상태, 행동에 대한 예측, 위험 수준 등 다양한 정보를 수집하고 메타데이터상에 기록하게 될 수 있다.

이러한 영상 분석 메타데이터는 직접적인 개인정보를 담고 있다고 볼 수 있다. 예를 들어 지능 치안 환경을 위해 용의자 등 특정 인물에 대한 CCTV 기반의 위치 추적 정보, 현재 행동, 위험도 판단 등 다양한 분석이 실시간으로 이루어질 수 있으며, 이러한 정보를 메타데이터로 저장할 경우 직접적으로 정적 개인정보뿐만 아니라 동적 개인정보까지 실시간으로 저장되므로 더욱 안전한 관리가 필요하게 될 것이며, 만약 메타데이터에 대한 해킹 등이 발생한 경우라도 반드시 안전이 보장되어야 할 필요가 있다[11-14].

공격자가 데이터베이스에 접근 권한을 획득하는 경우, 영상 메타정보에 대한 질의가 가능할 것이며, 이 과정에서 메타데이터를 평문으로 보관하고 있을 경우에는 해당 영상 메타 데이터가 공격자에게 그대로 노출될 것이다. 외부에서의 해킹 이외에 내부자에 의한 데이터 유출 공격 사례도 다수 존재하는 바, 개인정보보호 관점에서 메타데이터를 평문 그대로 저장하는 것은 심각한 개인정보 침해로 이어질 수 있어 매우 위험하다고 볼 수 있다. 따라서 영상 메타데이터에 적합한 보안 기술이 시급히 도입되어야 한다.

2.2.2 영상개인정보보호와 효율성의 비양립성

영상 메타데이터를 데이터베이스에 평문 형태로 저장하지 않고 암호화하여 저장할 경우, 데이터베이스에서의 질의가 매우 어렵게 된다는 문제점이 있다.

예를 들어, 영상 내의 특정 범위 내 위치에 접근한 사람을 검색하고자 할 경우, 암호화된 상태에서는 범위검색 질의를 할 수 없다. 이는 암호화된 데이터가 평문의 순서와 달라지는 것이 원인이며, 암호화된 데이터의 결과를 기준으로 범위검색을 수행한다면 원하는 결과와는 전혀 다른 데이터를 가져오게 된다.

이러한 문제를 해결하기 위한 방법으로 순서보존 암호화(OPE:Order-Preserving Encryption) 방식이 제안된 바 있다[15-17]. 그러나 이러한 OPE 방법을 사용하게 될 경우 암호화된 결과값이 평문과 순서가 동일하게 되므로 실질적으로 보안성에 있어 취약하다고 볼 수 있다. 또한, OPE 방식은 원본 데이터가 수치 데이터로 구성되어야 한다는 단점이 존재한다. 실질적으로 문자와 숫자의 정렬방식에는 차이가 존재한다. 데이터베이스에서는 예를 들어, 수치 데이터의 경우 100보다 20이 작은 값으로 간주되나, 문자열 데이터의 경우에는 20이 더 큰 값으로 간주된다. 영상메타데이터에는 다양한 종류가 있으며, 수치 데이터 또는 문자열 데이터로 취급된다. 기존의 순서보존 암호화 알고리즘은 수치 데이터의 처리만 가능하고, 문자열 데이터에 대한 처리는 불가능하여 영상 메타데이터의 보안기법으로는 적합하지 않다. 실질적으로 기밀성 확보와 데이터 검색에서의 효율성은 양립이 쉽지 않다. 원활한 데이터 검색을 위해서는 원본 데이터에 대한 최소한의 정보가 필요하여 실질적으로 기밀성 유지가 불가능하기 때문이다. 이와 같이 프라이버시 보호와 효율적인 영상 감시는 동시에 달성하기가 어려운 문제이다. 그러나, 향후 영상인식 기술이 점차 발전하고, 빅데이터 기반의 보다 세밀하게 분석된 영상정보가 메타데이터상에 상세히 기록될 것으로 예측되는 바, 영상감시 환경에서의 프라이버시 문제는 반드시 해결되어야 할 필요가 있다. 이미 법제도적으로 CCTV 영상 데이터에 대한 안전성 확보를 명시하고 있다. 현재 개인정보보호법 제25조(영상정보처리기기의 설치·운영 제한)의 제6항으로써 영상정보처리기기운영자는 개인정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 안전성 확보에 필요한 조치를 하여야 함을 명시하고 있어 이에 대한 고려가 반드시 필요한 상황이다.

III. 제안 방식

본 장에서는 본 논문에서 제안하는 영상 보안 감시환경 모델을 살펴보고, COP-변환 알고리즘과 이를 이용하여 지능형 영상감시 환경에서 CCTV 영상 데이터를 안전하게 보호할 수 있는 방법을 제안한다.

3.1 제안 방식 개요

3.1.1 영상 보안 감시환경 모델

Fig.3.은 본 논문에서 나타내는 영상 보안 감시환경을 나타낸다. 본 모델에서 CCTV는 실시간 영상을 촬영하여 클라우드 스토리지에 영상을 저장한다. 여기에서, 감시 시스템의 클라우드 스토리지에서는 해당 영상을 암호화하여 저장한다. 한편, CCTV 영상은 빅데이터 기반으로 실시간 메타정보가 추출되며, 메타정보는 본 논문에서 제안하는 COP-변환 알고리즘을 사용하여 저장한다. 이 과정에서, 데이터는 어떤 경우에도 평문으로 저장되지 않는다.

권한이 있는 CCTV 감시자는 CCTV 감시 시스템으로부터 영상 데이터를 수집하고 확인한다. 이 경우, COP-변환 메타데이터를 기반으로 비디오 영상에 대한 직접 질의가 가능하며, 권한이 있는 감시자는 이를 기반으로 원하는 영상을 검색하여 가져올 수 있다. 예를 들어 '오후 10시경 A 지역에 John 이 있는 영상은?' 이라는 방식의 질의를 통한 영상 검색이 가능하다. CCTV 메타정보는 일종의 개인정보로서, 평문으로 저장되면 보안상 큰 취약점이 존재한다. 따라서, 본 논문에서 제안한 COP-변환 기법을 이용하여 변환된 메타데이터로 저장한다.

이러한 경우, 평문 메타정보 및 평문 CCTV 영상 비디오 데이터는 CCTV 감시 환경에 남지 않게 되

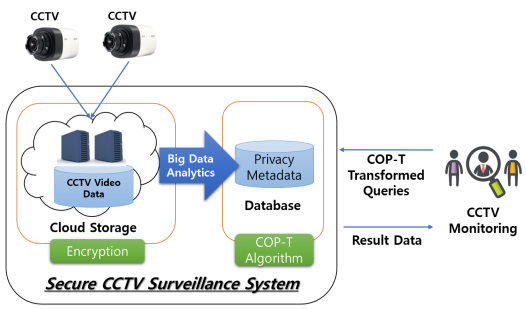


Fig. 3. Overview of Proposed Scheme

어, CCTV 영상데이터에 대한 프라이버시 보장이 가능하다. 즉, 공격자가 CCTV 감시 환경의 전산망에 해킹을 시도하여 비디오 영상정보 및 메타데이터의 탈취를 시도하여 실질적으로 영상정보 및 메타정보를 획득했다고 하더라도, 평문정보를 복원할 수 없다. 이러한 특성은 CCTV 감시 시스템 관련 운영업체 등 내부자에 의한 공격도 방지할 수 있다는 장점이 있다. 권한이 있는 감시자에 한해서 정상적인 방법으로 COP-변환 질의를 통하여 찾고자 하는 영상데이터를 정상적으로 가져올 수 있으며, 권한이 없는 접근자는 원본 데이터를 알아낼 수 없다.

3.2 세부 절차

3.2.1 약어

제안 기법 설명을 위한 약어는 Table 1.과 같다.

Table 1. Notation

Abbreviation	Content
D, R_{-1}	Pre-shared Initial Seed
C_i	i-th Character of C
P_i	i-th Random Scale Value
R_i	i-th Pseudorandom Number
$DIG(\cdot)$	Result of Digitalization
$CHR(\cdot)$	Result of Characterization
$PRNG(\cdot)^s$	Pseudorandom Number Generation
T_i	i-th COP-Transformation Value
x	The Last Element of Domain
n	Source String Length

3.2.2 COP 변환 알고리즘 설계

(1) COP 변환방식 개요

COP 변환방식은 원본 문자열을 변환식을 이용하여 변환된 문자로 치환하는 기법으로, 문자열을 구성하는 단일 문자 단위로 변환 문자로의 치환을 수행한다. 여기에는 다음과 같은 특징이 존재한다. 특정 문자열의 도메인과 변환된 문자열의 도메인은 각각 정렬 순서를 가지고 있다. 그러나 변환된 문자열 도메인의 정렬 순서는 평문과 반드시 동일한 순서를 따르지는 않으며, 아래와 같은 연산에 의해 결정된다.

$$C_i < C_z \rightarrow \begin{cases} \text{if } u \bmod 2 = 0 & T_i \le T_z \\ otherwise & T_i \ge T_z \end{cases}$$

Fig.4.는 COP 변환기법을 간단히 도식화하여 나타내며, Fig.5.는 실제 COP-변환 알고리즘을 나타내고 있다.

COP 변환기법은 단일 문자에 대한 수치화 단계, 의사난수를 기반으로 한 문자 도메인 기반 랜덤 측정

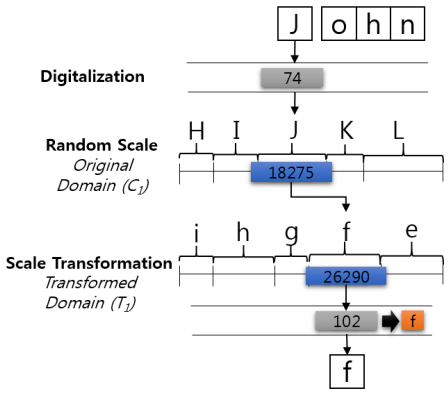


Fig. 4. The Concept of COP-Transformation

Algorithm : COP-Transformation

```
i=0
While  i≤n  do
  A_i = DIG(C_i)
  s=R_{i-1} + D
  R_i = PRNG(i)^s
  P_i = Σ_{k=1}^{A_i} PRNG(k)^s
  j=0, u=0
  while u≤P_i do
    M_j = PRNG(j)^{s+1}
    u = M_j + u
    j=j+1
  end
  if (R_i mod 2) = 0 then
    T_i = CHR(j)
  else
    T_i = CHR(DIG(x) - j)
  endif
  i=i+1
end
```

Fig. 5. COP-Transformation Algorithm

값을 연산하는 랜덤 스케일 단계, 랜덤 측정값을 새로운 측정값으로 변환하고 해당 스케일 변환값을 문자로 변환하는 스케일 변환 단계로 구성된다.

(2) 수치화(Digitalization) 단계

COP-변환기법은 원본 문자열에서의 단일 문자 단위로 치환 과정이 이루어진다. 여기에서, 먼저 단일 문자는 사전 정의된 매핑테이블에 근거하여 특정 숫자로 변환한다. 가장 간단한 방법으로, 아스키코드로 치환하는 경우를 생각해 볼 수 있다. 여기에서 아스키코드는 문자, 숫자, 특수문자를 모두 포함한다는 특성을 가지고 있다. 입력 문자에 따른 연산 결과 문자가 변환되어 출력되는 캐릭터 셋 영역은 매핑 테이블에서 정의되어 있는 전체 영역에서 발생할 수 있다. 따라서, 아스키코드를 매핑테이블로 정할 경우는 원본이 일반적인 숫자 및 문자만으로 구성되었다고 하더라도 COP-변환 결과값에는 특수문자가 포함될 것이다. 즉, 수치화 단계에서의 매핑테이블은 일반적인 아스키코드로 적용하더라도 연산에 문제는 없으나, 결과값에 특수문자를 포함하여 출력될 수 있다.

Table 6: ASCII Table

Char	Binary	Char	Binary	Char	Binary	Char	Binary
(nul)	00000000	(sp)	00100000	@	01000001	`	01100001
(soh)	00000001	!	00100001	A	01000010	a	01100010
(stx)	00000010	"	00100010	B	01000011	b	01100011
(etx)	00000011	#	00100011	C	01000100	c	01100100
(eot)	00000100	$	00100100	D	01000101	d	01100101
(enq)	00000101	%	00100101	E	01000110	e	01100110

Fig. 6. ASCII based Mapping Table[18]

(3) 랜덤 스케일(Random Scale) 단계

이 단계에서는 수치화 단계에서 생성된 수치에 대응하는 랜덤 스케일 값을 측정한다. 이를 위해 먼저 의사난수의 seed 값을 결정한다. 의사난수의 초기 seed는 사전 공유된 D값으로 정한다. D값은 원본 데이터를 COP-변환하는 과정에서 필요한 값이며, 향후 변환된 데이터에 대해 검색하는 경우에도 동일하게 해당 D값이 필요하며, D값을 알지 못하는 경우는 정상적인 검색을 수행할 수 없다. 최초의 단일 문자 변환 이후에는 R_i의 연산에 필요한 seed로써 R_{i-1}와 D를 더한 값을 취한다.

랜덤 스케일값 P_i는 의사난수인 $PRNG(k)^s$의 합에 의해 결정된다. 원본 문자열에서의 i번째 문자에 대응하는 랜덤 스케일값 P_i는 아래의 식으로 구할 수 있다.

$$P_i = \sum_{k=1}^{A_i} PRNG(k)^s$$

(4) 스케일 변환(Scale Transformation) 단계

이 단계는 앞서 랜덤 스케일 단계에서 생성한 P_i 값을 기반으로 새로운 매핑 숫자값을 연산하는 단계이다. 여기에서는 앞서 연산한 seed 값인 s에서 1을 더한 결과를 이 단계에서의 seed로 정한다. 이후, 해당 seed를 기반으로 발생된 의사난수의 전체 합계가 랜덤 스케일 단계의 결과값인 P_i에 이르기까지 반복하여 생성하고, 해당 횟수 j를 카운팅한다. 이러한 부분은 Fig.5.에서 $PRNG(j)^{s+1}$을 기반으로 u를 생성하는 과정에서 나타나 있다.

이 결과에 의해 연산된 반복 횟수 j는 최종 결과값의 문자를 얻기 위한 매핑테이블에 대한 입력값이 될 것이다. 스케일 변환에서의 seed는 앞서 언급한 랜덤 스케일 단계에서의 seed와 다르게 적용되었으므로 최초 수치화된 결과값과 여기서 발생된 결과값은 서로 상이하다. 이 결과값을 앞서 수치화 단계에서의 매핑테이블을 역으로 적용하여 특정 문자로 치환하는 것으로 최종 변환된 결과값을 구할 수 있다. 앞서 언급한 단계를 문자열이 끝날때까지 반복하여, 전체 문자열에 대응하는 COP-변환 문자열을 구할 수 있다.

3.2.3 CCTV 영상 데이터 매핑 구조

CCTV 영상은 스토리지에 암호화되어 저장된다. 암호화 알고리즘으로 AES 등 대칭키 암호화 알고리즘이 사용될 수 있으며, 암호화된 비디오 파일은 파일 단위로 Video ID가 부여되고, 하나의 비디오 파일은 세부적으로 여러 파일로 분할하여 각각 Chunk ID를 가지게 된다. CCTV 영상 파일은 대용량이라는 특성을 가지고 있어, 특정 비디오 영상 파일 내에는 별도로 다수의 Chunk 영역을 분할하여 가지고 있어야 한다. 예를 들어, John이 출현한 CCTV 영상파일 목록을 찾고자 할때, 특정 날짜에 촬영된 전체 파일을 복호화하여 찾게 된다면 효율성을 크게 떨어뜨릴 수 있다. 이 경우, 각 영상 파일에 John이 나타난 세부적인 Chunk 목록을 확보할 수 있다면, 해당 부분 파일에 대한 복호화만 수행하면

되므로 영상 복호화의 효율성에 있어 큰 이점을 가질 수 있다. 따라서, 본 논문에서는 영상 데이터를 세부적인 Chunk 파일로 분할하여 별도로 암호화하여 보관하는 구조를 제안한다. 이 경우, Tag 정보에 대한 검색에 따라, 해당 정보에 매칭되는 Video ID 및 Chunk ID를 확보하여, 해당 Chunk ID에 대응하는 파일의 일부만 복호화를 할 수 있어 성능상 효율성을 가져올 수 있다. 이 과정에서 CCTV 비디오 파일만 암호화가 되어 있고, 메타데이터를 평문으로 저장할 경우는 메타데이터 그 자체만으로 비디오 파일의 내용을 상세히 노출하고 있으므로 프라이버시 보호에 큰 문제가 될 수 있다. 따라서, 본 논문에서는 CCTV 영상을 암호화함과 동시에, 영상메타도 COP-변환값으로 저장하여 메타DB 및 스토리지상 어느곳에도 평문을 저장하지 않으므로 영상객체의 개인정보를 보호할 수 있다. Fig.7.은 영상 메타데이터와 실제 영상 데이터 간의 매핑 구조를 나타낸다.

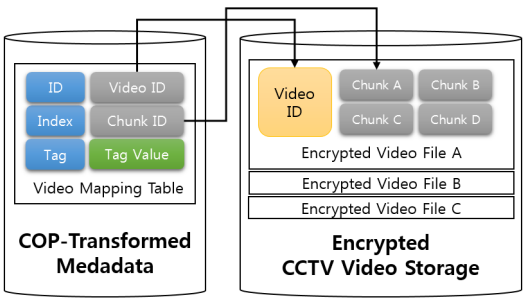


Fig. 7. Video Data Mapping Structure

3.2.4 COP 변환메타 질의 기법

COP-변환기법으로 변환된 메타값은 데이터베이스상에 평문이 아닌 변환메타로 입력된 상태에서도 메타데이터에 대한 직접 질의가 가능하다는 장점이 있다. COP-변환 데이터는 직접적으로 전방일치 및 범위검색 뿐만 아니라 통계를 위한 집계검색이 가능하다는 특징이 있다. 또한, 데이터베이스 인덱스 구성을 그대로 활용할 수 있다는 장점이 있다. 실질적으로 암호화된 데이터베이스는 일치검색 질의 이외에의 전방일치, 범위검색 등에 인덱스 사용이 불가능하여 데이터베이스의 효율을 크게 저하시키는 원인이 된다. 그러나 COP-변환 기법은 데이터베이스 질의시 인덱스 이용이 가능하여 평문 상태에서 질의하는 것과 효율성에서 차이가 없다고 볼 수 있으므로, 평문 데이터를 노출하지 않으면서 DB 질의 과정에서도 성능의 향상을 가져올 수 있다는 큰 장점이 있다.

Fig.8.은 통상적인 SQL 범위검색 질의를 COP-변환 질의로 변경하는 방법을 나타낸다. COP 변환값은 평문의 순서를 정순, 혹은 역순으로 랜덤하게 저장하고 있으며, 이는 COP 변환 과정에서의 스케일 변환(Scale Transformation) 단계의 결과값인 u 값을 기준으로 결정한다. 즉, R_i 값의 mod 연산에 대한 결과값에 따라 범위검색시 대소 구분이 달라진다는 특징이 있다. COP-변환 데이터는 평문의 정렬순서가 문자열의 자릿수 단위로 정순 혹은 역순으로 랜덤하게 적용된다는 특징을 가지고 있으며, R_i값을 알 수 있는 경우에만 데이터에 대한 정상적인 질의가 가능하다. 여기에서, 범위검색, 집계검색 등은 평문의 순서가 역순이더라도 조건값을 변경하면 질의가 가능해진다는 특징이 있다. 예를 들어, 상위 30%의 데이터를 가져오고자 하는 경우, 역순에서는 하위 30%의 데이터를 가져오는 것으로 동일한 결과를 출력할 수 있다. 따라서, 제안한 COP-변환 방식을 적용하면 전체 순서를 정순으로 유지하지 않으면서도 범위/집계검색 처리가 가능하다는 특징이 있다.

만약, 적합한 권한을 가지고 있지 않은 해커 등 공격자의 경우는 사전 공유된 D, R_{-1}값 및 의사난수의 seed값인 s를 알지 못하므로 C_a 및 C_z의 값을 알고 있다 하더라도 T_a 및 T_z 값을 생성할 수 없으므로 변환 쿼리를 구성할 수 없다.

Original Query :

select video_id, video_idx
from metadata
where tag $\rangle$ C_a *and tag* $\langle$ C_z

Transformed Query :

select video_id, video_idx
from metadata

$$\begin{cases} where\ tag \geq T_a \text{ and } tag \leq T_z, & \text{if } R_i \bmod 2 = 0 \\ where\ tag \leq T_a \text{ and } tag \geq T_z, & otherwise \end{cases}$$

Fig. 8. COP-Transformation of SQL Query

IV. 구 현

4.1 성능 측정

본 절에서는 제안 방식의 성능 측정 결과를 살펴본다. COP-변환 알고리즘에 대한 성능 측정 환경은 다음과 같다. 알고리즘은 C++로 구현하였으며, 데이터베이스는 SQLite를 사용하였다. CPU는 i7-4790K@4.0GHZ 환경에서 수행하였으며, 메모리 사이즈는 6GB에서 수행하였다. 측정 결과는 Fig.9.와 같다. 평문으로 데이터베이스 질의를 수행하였을 때가 가장 성능이 높게 나오지만, 실질적으로는 순서유지 암호화 방식 및 제안하는 방식과 비교하면 큰 차이점은 존재하지 않는다. 이는 세 방식 모두 데이터베이스의 인덱스를 활용함으로써 빠른 속도로 데이터를 가져올 수 있으며, 인덱스된 데이터베이스 처리시간에 암호화 처리 시간만 추가되기 때문이다. 한편, AES 암호화를 수행하였을 경우는 데이터베이스 인덱스를 지원하지 않으므로 데이터베이스 질의에 매우 큰 오버헤드가 소요되는 것을 확인할 수 있다.

평문으로 데이터베이스에 질의할 경우와, 본 논문에서 제안한 COP-변환 기법으로 질의할 경우에는 질의 처리 속도 차이가 미미한 수준임을 알 수 있다. 기존의 순서유지 암호화 방식에서도 이러한 측면에서는 동일한 특성을 가지고 있으나, 기존 순서유지 암호화 방식은 평문과 동일한 순서를 가지고 있어 데이터에 대한 분석 공격이 가능함으로써 보안상 취약점이 존재한다. 본 논문에서 제안한 COP-변환 기법은 평문 및 순서유지 암호화 방식의 수준과 대등한 성능을 보이고 있으며, 실질적으로 순서 정보가 평문과 동일하게 적용되지 않고, 역순, 정순이 랜덤하게 적용되어 더욱 안전성이 높다는 장점이 있다. 한편, AES와 같은 일반적인 암호화 알고리즘의 경우는 데이터베이스의 인덱스 사용이 불가능하여 매우 큰 오버헤드를 발생시키며, 기존의 암호화 방법에 비해 제안한 방식이 더욱 효율적임을 나타내고 있다.

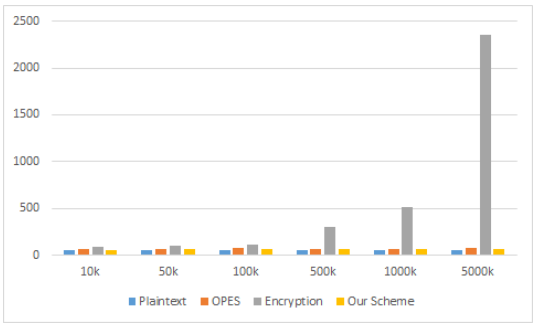
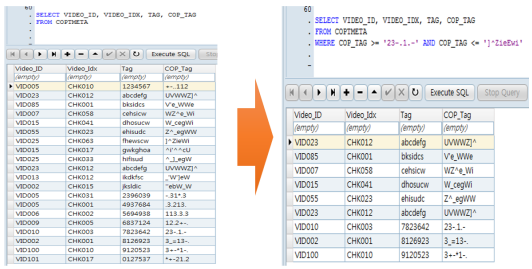


Fig. 9. The time overhead in the range query

4.2 영상메타 질의처리

Fig.10.에서의 왼쪽 그림은 전체 데이터를 나타내고 있으며, 오른쪽 그림은 COP-변환을 통하여 재구성한 범위검색 질의를 나타낸다. COP-변환 질의 출력은 평문으로 구성된 데이터베이스 환경에서의 평문 질의 결과와 동일한 결과를 출력하는 것을 알 수 있다. 또한, 평문과 동일한 실행계획(Explain Query Plan)을 가지므로 변환된 상태에서도 인덱스를 평문과 동일하게 그대로 활용할 수 있어 질의상 효율적인 성능을 보인다.



(a) Total Data (b) Range Data

Fig. 10. COP-Transformation of SQL Query

V. 분 석

5.1 안전성 분석

5.1.1 영상 데이터의 외부 노출

영상 데이터는 메타정보가 저장된 메타 데이터베이스와 실제 CCTV 영상이 저장된 영상 스토리지로 구성되어 있다. CCTV 영상 데이터는 그 자체만으로 개인정보의 성격을 가지고 있다. 비디오 영상에 촬영된 개인의 위치, 동선을 CCTV 영상정보만으로 직접적으로 파악할 수 있기 때문이다. 이러한 분석 정보는 영상 메타데이터상에 기록되며, 영상 CCTV 데이터 및 영상 메타 데이터는 그 자체로서 프라이버시 침해 요소를 지니고 있다고 볼 수 있다. 이에 대한 가장 안전한 방법은 데이터를 암호화하는 것이다. 본 논문에서 제안하는 방법에서는 실제 스토리지 상

에 저장되어 있는 CCTV 영상 데이터는 모두 암호화를 적용하므로 공격자에게 스토리지상의 영상데이터가 노출되더라도 안전하다. 한편, 영상 메타데이터는 COP-변환 방식으로 변환된 데이터로 구성되어 있으므로 공격자는 초기 비밀 공유값인 D와 의사난수 발생에 필요한 초기 seed를 알지 못하면 COP 변환 값을 구성할 수 없다. 즉, 메타정보, CCTV 영상정보 어디에도 직접적으로 개인정보를 평문 형태로 저장하지 않으므로 불가피하게 영상데이터 전체가 외부에 노출되더라도 안전성을 보장할 수 있다.

5.1.2 메타 질의과정 노출

검색을 위한 메타정보 질의시 메타정보 검색에 사용된 SQL문 및 질의 실행 결과값을 획득하는 것만으로 영상 CCTV 내용의 유추가 가능할 수 있다. 그러나, 제안한 방식에서는 영상메타 질의시 SQL상에 평문 메타를 노출하지 않는다. 또한, 데이터베이스상의 메타정보값 전체가 COP-변환 기법으로 변환되어 있으므로 SQL 질의를 통하여 검색된 결과값 또한 변환값 형태로 리턴한다. 따라서, 영상메타 질의 및 리턴 데이터를 공격자가 모두 확보하게 되더라도, 해당 변환값으로부터 원본 평문을 복원할 수 없어 안전하다.

5.1.3 내부자 공격

CCTV 비디오 영상은 대용량 데이터이므로 클라우드 스토리지상에 보관하는 경우가 많다. 특히, 클라우드 환경에서는 클라우드 시스템 운영자 등 내부자 공격에 주의할 필요가 있다. 여기에서, 내부자는 데이터베이스와 스토리지에 상시 접근이 가능하다. 본 논문에서 제안하는 기법은 초기 공유값인 D를 인가된 권한이 있는 자에게만 공유한다. 또한 메타정보를 포함한 데이터베이스는 COP 변환 기법을 적용하여 데이터가 변환되어 있으며, 스토리지 CCTV 영상 데이터는 전체 파일이 암호화되어 있다. 여기에서, 내부자 등 관리자 입장에서의 파일 및 데이터 접근 권한과 실제 영상감시 처리 업무자의 평문 데이터 확인 권한은 서로 다르게 적용될 필요가 있다. 즉, 초기 공유값인 D와 R_{-1}을 실질적인 정보 열람 권한이 있는 자에게만 공유하여야 한다. 따라서, 클라우드 스토리지 관리자는 암호화된 상태의 데이터에 접근은 가능하나, 정보 열람에 대한 권한이 없으면 데이터베이스 및 스토리지로부터 평문 정보를 복원할 수 없다.

5.2 효율성 분석

5.2.1 영상 데이터 복호화 성능 관점

CCTV 영상 데이터 전체가 암호화된 경우, 이에 대한 메타 질의 등을 수행하더라도, 특정한 조건에 맞는 영상 파일을 획득하려면 해당 파일 전체를 복호화하여야 한다. 이러한 점은 데이터 검색 속도를 현저히 떨어뜨리는 요인이 될 수 있어 비효율적이다. 제안하는 방식은 데이터 암호화 시 Chunk 단위로 파일을 여러 단위로 분할하여 암호화를 하는 특징을 가지고 있으며, 영상메타 데이터에 대한 질의 수행시 영상 파일의 일부 조각에 해당하는 부분 영상 파일에 대한 Chunk ID를 기준으로 처리할 수 있다. 따라서, 제안하는 방식은 전체 CCTV 영상데이터를 복호화하지 않고 Chunk 단위의 검색 조건에 부합하는 부분영상정보만을 획득하여 복호화를 수행하므로 데이터 복호화 성능 관점에서 효율성을 갖는다.

5.2.2 질의문 구성의 효율성

메타정보 검색을 위한 질의문 구성시 COP-변환 기법을 적용한 데이터는 평문과 동일한 수준의 질의 처리 효율성을 갖는다. 즉, COP-변환 데이터베이스는 일치검색, 전방일치검색, 범위검색 뿐만 아니라 복수의 테이블에 대한 JOIN을 활용한 질의도 가능하다. 또한, 메타정보에 대한 통계분석 시에도 질의문을 간단히 구성할 수 있다. 예를 들어, 특정 조건하에서의 MIN, MAX, COUNT, AVG와 같은 통계에 필요한 집계(Aggregation) 질의를 COP-변환 메타데이터상에 그대로 적용할 수 있다는 장점이 있다. 이러한 장점은 데이터베이스를 단순 암호화한 상태에서는 달성할 수 없다. 만약, 데이터베이스를 일반적인 암호화 방식으로 암호화하게 되는 경우 실질적으로 일치검색 이외에는 질의문 구성이 매우 어렵게 된다. 예를 들어, 범위검색 질의가 필요한 경우, 암호화된 데이터베이스는 평문의 결과값과 정렬 순서가 서로 상이하므로 전체 데이터를 복호화한 데이터를 별도로 보관하고 있지 않는다면 직접적인 SQL 범위검색 질의 구성은 불가능하다. 본 논문에서 제안하는 방식은 COP-변환 메타 그대로 범위검

색 적용이 가능할 뿐만 아니라, 데이터베이스 인덱스 정보를 그대로 활용할 수 있어 속도 측면에서 크게 효율적이다. Table 2.는 메타정보에 대한 평문, 암호화, COP-변환의 경우에 대한 질의 처리 가능 여부를 나타내고 있다. 평문의 경우는 일치검색, 범위검색, 집계검색이 가능하나, 메타데이터를 암호화할 경우는 일치검색 이외 범위검색, 집계검색에 대한 질의문 구성이 불가능하다. 제안하는 기법은 평문과 동일하게 일치검색, 범위검색, 집계검색을 지원하며, 데이터베이스의 인덱스를 그대로 활용 가능하여 효율적이다.

Table 2. Efficiency Comparision

Query Type	Plaintext	Encryption	The Proposed Method
Equation Query	○	○	○
Range Query	○	×	○
Aggregation Query	○	×	○

VI. 결 론

향후 인공지능 기술의 발달에 따라, CCTV 기반 지능형 영상분석 기술 및 시장이 크게 발전하게 될 것이다. 그러나, CCTV 촬영 영상은 개인정보를 그대로 가지고 있으므로, 이에 대한 정보보호 대책이 필수적이다. 기존의 영상데이터 보호 방식은 영상 마스킹, 혹은 단순 데이터 암호화 방식에 의존하고 있으며, 영상 메타 데이터 기반의 효율적이고 안전한 CCTV 영상 검색 기능은 제공해 주지 않는다. 실질적으로 메타데이터를 암호화하게 되면 일치검색을 제외한 검색이 불가능하여 메타정보는 평문으로 저장하는 경우가 일반적이다. 그러나 메타정보 자체가 영상에 대한 많은 정보를 포함하게 될 것이며, 향후 빅데이터 기반 영상 분석 기술이 발전할수록 영상 메타데이터상에 더욱 많은 개인정보를 노출하게 될 것이다.

따라서, 본 논문에서는 COP-변환 기법을 제안하였다. 이 방식은 평문을 노출하지 않은 상태에서 데이터베이스 질의를 평문과 동일하게 수행 가능하여 영상 메타데이터의 효율을 크게 높일 수 있다는 장점이 있다. 제안한 방식의 설명을 위해 먼저 2장에서 기존의 영상정보보호 기법과 영상메타데이터 보호의 필요성에 대하여 살펴보았고, 3장에서는 본 논문에서 제안한 COP-변환 알고리즘과 영상 메타정보 검색을 위한 질의문 구성 방법을 설명하였다. 이후 4장에서는 제안한 기법을 실제 환경에 구성하여 성능 측정을 수행하였으며, 5장에서는 안전성과 효율성 관점에서 제안한 알고리즘을 분석하였다.

향후 빅데이터 기반의 보다 세밀한 영상 분석이 가능해질 것으로 예상됨에 따라, 앞으로의 영상 메타데이터는 더욱 많은 정보를 포함하게 될 것이다. 이는 CCTV 영상 피촬영자의 프라이버시 침해와 직접적으로 연결될 수 있으므로 이에 대한 안전한 대책이 필수적으로 선행되어야 한다. 앞으로 4차산업혁명의 시대가 도래함에 따라, 지능형 영상 감시 환경의 보안대책에 대한 더욱 많은 연구가 필요할 것이며, 안전한 미래사회를 위한 필수적인 프라이버시 보호 기술로 자리잡게 될 것으로 보인다.

References

[1] Namje Park, Marie Kim, "Implementation of load management application system using smart grid privacy policy in energy management service environment", Cluster Computing, 17(3), pp. 653-664, Sep. 2014.

[2] Donghyeok Lee, Namje Park, "Geocasting-based synchronization of Almanac on the maritime cloud for distributed smart surveillance", The Journal of Supercomputing, pp.1-16, 2016.

[3] Namje Park, Hongxin Hu, Qun Jin, "Security and Privacy Mechanisms for Sensor Middleware and Application in Internet of Things (IoT)", International Journal of Distributed Sensor Networks, 2016, 2016.

[4] F. Dufaux, T. Ebrahimi, "Scrambling for Privacy Protection in Video Surveillance Systems," IEEE Trans. on Circuits and Systems for Video Technology, 18(8), pp. 1168-1174, 2008.

[5] P. Agrawal, P.J. Narayanan, "Person De-identification in Videos," IEEE

Trans. on Circuits and Systems for Video Technology. 21(3), pp. 299-310, 2011.

[6] F. Dufaux, T. Ebrahimi, "A Framework for the Validation of Privacy Protection Solutions in Video Surveillance," IEEE International Conference on Multimedia and Expo (ICME), pp. 66-71, 2010.

[7] E. M. Newton, et. al., "Preserving Privacy by De-identifyuing Face Images," IEEE Trans. on Knowledge and Data Engineering, 17(2), pp. 232-243, 2005.

[8] Cha Gun Sang, Shin Yong Tae, "Personal Video Privacy Issue to Increasing CCTV Installation," Journal of Computing Science and Engineering, 27(12), 2009.

[9] F. Peng, X. Zhu, M. Long, "A ROI Privacy Protection Scheme for H.264 Video Based on FMO and Chaos," P. Agrawal, P.J. Narayanan, IEEE Trans. on Information Forensics and Security, 8(10), pp.1688-1699, 2013.

[10] D. A. Rodríguez-Silva, L. Adkinson-Orellana, F. J. González-Castaño, I. Armiño-Franco, "Video surveillance based on cloud storage", 2012 IEEE Fifth International Conference on Cloud Computing, pp. 991-992, 2012.

[11] Lee D, Park N, "Electronic identity information hiding methods using a secret sharing scheme in multimedia-centric internet of things environment", Personal and Ubiquitous Computing, DOI 101007/s00779-017-1017-1, 2017.

[12] Donghyeok Lee, Namje Park, "A Proposal of SH-Tree Based Data Synchronization Method for Secure Maritime Cloud", Journal of the Korea Institute of Information Security & Cryptology, 26(4), pp. 929-940, Aug. 2016.

[13] Donghyeok Lee, Namje Park, "A Proposal of SH-Tree Based Data Synchronization Method for Secure Maritime Cloud", Journal of the Korea Institute of Information Security & Cryptology, 26(4), pp. 929-940, Aug. 2016.

[14] Namje Park and Namhi Kang, "Mutual Authentication Scheme in Secure Internet of Things Technology for Comfortable Lifestyle", Sensors, 16(1), pp. 1-16, Dec. 2015.

[15] Agrawal, Rakesh, et al. "Order preserving encryption for numeric data." Proceedings of the 2004 ACM SIGMOD international conference on Management of data. ACM, 2004.

[16] Dongkook Kim, Hyeok Lee, "Personal Information De-Idenification Trends based on Big Data", Review of Korean Society for Internet Information, 16(2), pp.15-22, Dec. 2015.

[17] Donghyeok Lee, Namje Park, "A Study on Metering Data De-identification Method for Smart Grid Privacy Protection",Korea Institute of Information Security and Cryptology, 26(6), pp.1593-1603, Dec. 2016.

[18] Aiden A. Bruen and Mario A. Forcinito, "Cryptography, Information Theory, and Error-Correction: A Handbook for the 21st Century", John Wiley & Sons, Inc., 2005

〈저자소개〉

이 동 혁 (Donghyeok Lee) 정회원
2007년 2월:동국대학교 전자상거래기술전공 공학석사
2018년 2월:제주대학교 컴퓨터교육전공 공학박사
2007년 6월~2008년 5월:한국전자통신연구원 정보보호연구단 연구원
2008년 11월~2015년 6월:KT 플랫폼개발단 과장
2015년 9월~현재:제주대학교 초등교육연구소 특별연구원
2018년 8월~현재:제주대학교 과학기술사회(STS)연구센터 학술연구교수
〈관심분야〉 클라우드 보안, 스마트그리드 보안, 데이터베이스 보안, 해사클라우드 등

박 남 제 (Namje Park) 종신회원
2008년 2월:성균관대학교 컴퓨터공학과 박사
2003년 4월~2008년 12월:한국전자통신연구원 정보보호연구단 선임연구원
2009년 1월~2009년 12월:미국 UCLA대학교 공과대학 Post-Doc.,WINMEC 연구센터 Staff Researcher
2010년 1월~2010년 8월:미국 아리조나(ASU) 주립대학교 컴퓨터공학과 연구원
2010년 9월~현재:제주대학교 초등컴퓨터교육전공 교수, 과학기술사회(STS)연구센터장
2018년 3월~현재:제주대학교 일반대학원 융합정보보안전공 학과장
〈관심분야〉 융합기술보안, 컴퓨터교육, 스마트그리드, IoT, 해사클라우드 등