논문 2023-4-11 http://dx.doi.org/10.29056/jsav.2023.12.11

영지식 증명 기반 OTT 플레이어 에이전트 환경의 콘텐츠 시청 정보 비식별화 및 검증 방법에 관한 연구

박병찬*, 장세영*, 김석윤*, 김영모*†

Proposition of Viewing Information De-identification and Verification Method in OTT Player Agent Environment Based on Zero-Knowledge Proof

Byeongchan Park*, Seyoung Jang*, Seok-Yoon Kim*, Youngmo Kim*†

요 약

코로나 팬데믹 현상으로 인하여 '넷플릭스'와 같은 OTT(Over The Top) 서비스를 제공하는 다양한 영상 제공 플랫폼 등이 급격한 성장을 이루었다. 이러한 영상 콘텐츠를 제공하는 OTT 플랫폼에서는 개인정보보호 등의 여러 이슈로 이용자가 시청하는 콘텐츠의 시청 시간을 공개하지 않고 있어 이해관계자들은 제3자 방식의 이용행태 조사를 하고 있는 실정이다. 객관적이고 신뢰성 있는 제3자 방식의 이용행태조사 방법이 필요한 한편, 그 과정에서 개인정보 보호에 대한 노력도 필요하다. 이용행태조사를 위해 시청률 조사에서 필요한 정보는 어떤 콘텐츠가 얼마나 이용되는 지에 대한 정보만 필요하며, 이용자에 대한 개인정보는 필요하지 않다.

본 논문에서는 이용자가 OTT 서비스를 제공하는 OTT 플레이어 에이전트를 통해 특정 OTT 플랫폼을 선택하여 콘텐츠를 이용하였을 때 생성되는 이용 정보를 영지식 증명 프로토콜을 이용하여 비식별화로 생성 및 검증할 수 있는 방법을 제안한다.

Abstract

Due to the Corona-19 pandemic phenomenon, various video service platforms that provide OTT (Over The Top) services such as 'Netflix' have achieved rapid growth. OTT platforms that provide such video content do not disclose the viewing time of the content watched by users due to various issues such as personal information protection, so stakeholders are conducting third-party research on usage patterns. While an objective and reliable third-party method of investigating usage patterns is needed, efforts to protect personal information are also needed in the process. The information required for viewership surveys to investigate usage patterns only requires information about what content is used and to what extent, and does not require personal information about users.

In this paper, we propose a method to generate and verify the generated usage information by de-identification using zero-knowledge proof protocol when a user selects a specific OTT platform and uses content through an OTT player agent that provides OTT services.

한글키워드 : OTT, 시청률 조사, 영지식 증명, 비식별화

keywords : OTT, Viewer Rating Survey, Zero-Knowledge Proof, De-identification

* 숭실대학교 컴퓨터학과

† 교신저자: 김영모(email: ymkim828@ssu.ac.kr)

접수일자: 2023.12.09. 심사완료: 2023.12.15.

게재확정: 2023.12.20.

1. 서 론

코로나 팬데믹 현상으로 인하여 다양한 비대면 서비스가 짧은 시간에 많은 성장을 이루었다[1]. 대표적으로 '넷플릭스'와 같은 OTT(Over The Top) 서비스를 제공하는 다양한 영상 제공 플랫폼 등이 있다[2]. 2022년 5월 한국정보통신정책연구원에서 발간한 보고서에 의하면 2019년 41.0%, 2020년 72.2%, 2021년 81.7%로 OTT 서비스 이용률이 급격하게 증가하였다고 발표하였다[3]. 이러한 OTT 서비스를 제공하는 플랫폼에서는 많은 콘텐츠를 유통하고 있다. 하지만 OTT 플랫폼에서는 개인정보보호 등의 여러 이슈로 이용자가 시청하는 콘텐츠의 시청 시간을 공개하지 않고 있어 이해관계자들은 제3자 방식의 이용행태 조사를 하고 있는 실정이다[4]. 객관적이고 신뢰성 있는 제3자 방식의 이용행태조사 방법이 필요 하다. 그러나 개인정보 보호에 대한 노력도 필요하다. 이용행태조사를 위해 시청률 조사에서 필요한 정보는 어떤 콘텐츠가 얼마나 이용되었는지만 필요하며, 이용자 개인을 특정하는 개인정보는 필요하지 않다[5]. 2023년 9월에 개정된 '개인정보보호법 시행령' 제4장 '개인정보의 처리'의 제 14조2의 제1항에 의하면, "가명처리 또는 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부"에 대한 내용이 법으로 명시되어있다.

본 논문에서는 이용자가 OTT 서비스를 제공하는 OTT 플레이어 에이전트를 통해 특정 OTT 플랫폼을 선택하여 콘텐츠를 이용하였을 때 생성되는 이용 정보를 영지식 증명 프로토콜을 이용하여 비식별화로 생성 및 검증할 수 있는 방법을 제안한다.

본 논문의 구성으로 2절에서는 관련연구로 가명처리 기법 및 영지식 증명에 대해 기술한다. 3절에서는 본 논문에서 제안하는 영지식 증명 기반 시청 정보 비식별화 및 검증 방법에 대해 기술하며, 4절에서 결론으로 마무리한다.

2. 시청 정보 생성 데이터 및 영지식 증명

2.1 OTT 기반 시청 정보 생성 및 비식별화 범위

TV, IP TV 등에서 시행되는 시청 조사는 이용자 성별 및 나이 등의 포함된 개인정보 및 이용되는 콘텐츠 정보를 포함하여 최종 시청 정보를 생성하게 된다[6]. 이를 바탕으로 최종 시청률을 산출하게 된다. 언론에 노출되는 시청률 정보는 지역별, 연령별, 성별 등으로 어느 콘텐츠가 얼마나 사용되었는지를 보여준다. 이러한 정보는 인물을 특정하지 않아도 현재 가장 인기있는 콘텐츠가 어떤 콘텐츠인지를 알 수 있다[7]. 하지만 인터넷 환경에서는 개인정보의 암호화 등 다른 특정한 조치를 취하지 않는 이상 누군지 바로 특정할 수 있다. 예를들어 인터넷 방송 같은 경우 내가 어떤 방송을 선택하였고, 얼마나 시청하였는지를 바로 산출할 수 있다. 콘텐츠의 권리자 또는 권리 단체는 특정 인물을 알 필요 없기 때문에 이러한 개인정보를 비식별화하여야 한다[8].

국내·외 OTT 서비스 플랫폼에서 회원 가입시 개인정보 수집 범위는 ID 대용으로 쓰이는 E-mail 속성, 성인 콘텐츠로 인한 나이 속성 그리고 이름 대신 플랫폼에서 쓰이는 프로필을 최소로 다양한 개인정보를 수집한다. 이러한 이용자 정보는 권리자 등 이해관계자들에게 제공할 수 없기 때문에 이용자 정보는 비식별화하며, 그림 1과 같다.

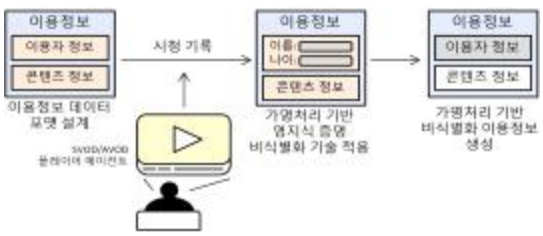

그림 1. 시청 정보 비식별화 범위
Fig. 1. De-identification range of viewing information

2.2 영지식 증명기반 가명 정보 처리

영지식 증명 기반 데이터의 비식별화를 위해 식별자(Identifiers)와 속성자(Attribute value)를 활용한다[9]. 식별자란 개인을 특정할 수 있는 고유 식별 정보와 상세 주소 등의 정보이며 속성자란 해당 데이터 자체는 식별정보가 아니지만 다른 정보들을 통하여 정보 주체를 추론할 수 있는 데이터이다.

데이터 비식별화 방법 중 가명 정보 처리 방법은 휴리스틱 가명화와 암호화 방법으로 구분 되며 이 중 암호화 방법의 대표적 처리 방법으로 동형암호(Homomorphic Encryption)와 영지식 증명(Zero-Knowledge Proof), 차등 정보 보호(Differential Privacy),다자간 계산(Secure Multiparty Computation)을 들 수 있다[10-12].

이중 영지식 증명은 증명자(Prover)가 자신이 가지고 있는 비밀에 대한 정보를 노출하지 않고 자신이 비밀스러운 정보를 가지고 있다는 것을 확인자(verifier)에게 증명하는 기술이다. 최근에는 자신이 가지고 있는 비밀스러운 정보의 범위를 증명하는 range proof 방식을 적용하도록 연산 처리되어 고도의 익명성을 요구하는 데이터의 비식별화에 많이 활용되고 있다.

3. 영지식 증명 기반 데이터 비식별화 및 외부 공유를 위한 재구성된 자격 증명 방법

3.1 영지식 증명 기반 비식별화 및 외부 공유를 위한 구성

OTT 플레이어 에이전트 환경에서 시청 정보 비식별화 및 데이터 공유 위한 구성은 증명인-발행인-검증인-외부인으로 구성되며, 그림 2와 같다.

그림 2. 비식별화 및 검증을 위한 구성
Fig. 2. Overall structure for de-identification and verification

증명인은 시청 정보를 실제 보유한 주체로 시청 정보에 대한 증명, 정보의 거래 등 최종 권한을 가진다. 발행인은 증명인이 주장하는 시청 정보를 실제 자격 증명으로 발행하는 역할로 VC를 발행한다. 또한, 발행된 원본 VC 외 재구성된 자격 증명인 VC_R을 발행할 수 있는 권한을 지닌다. 그리고 검증인이 재구성된 자격 증명을 외부에 공유할 때, 재구성된 자격 증명이 자신의 의해 발행되었음을 입증할 수 있어야 한다. 검증인은 증명인의 VC를 검증하며, 영지식 증명 기반 신원 속성의 결격 여부를 검증한다. 또한, VC에 대한 검증만 수행하는 역할이 아닌 재구성된 자격 증명인 VC_R을 발행할 수 있는 개체가 된다. 이에 대한 데이터 비식별화 및 데이터 공유 방법은 그림 3과 같다.

그림 3. 데이터 비식별화 공유 방법
Fig. 3. Sharing method of data de-identification

외부인은 검증인이 재구성한 자격증명인 VC_R를 제공받는 개체이다.

이러한 증명인-발행인-검증인-외부인은 OTT 플레이어 에이전트 환경에서 시청 정보 비식별화 및 검증을 위해 다시 구성하면, 표, 1과 같다.

표 1. OTT 환경 비식별화 및 검증을 위한 구성
Table 1. Construction for de-identification and verification in OTT environment

기존 구성	재구성
증명인	시청자
발행인	OTT 플레이어 에이전트
검증인	시청조사단체
외부인	권리자 등

외부인은 권리자 등으로 자신의 콘텐츠가 얼마나 시청되었는지 확인할 수 있다.

3.2 외부 공유를 위한 재구성된 시청 정보의 자격 증명

VC가 검증될 경우 ZKP의 검증 결과가 발생된다. 예를들어 시청자가 이용하는 콘텐츠의 시청 정보는 이용자 정보 및 콘텐츠 이용 정보로 나눌 수 있는데, 이 중 이용자 정보는 이름, 나이, 성별 등으로 그 중 콘텐츠 시청 가능 연령이 19세 이상일 경우 시청자의 정확한 나이 정보가 아닌 나이 속성만을 가지고 시청 가능한지 참, 거짓의 정보만을 획득할 수 있다. 이러한 특성으로 인하여 정확한 속성 정보가 필요로 하지 않더라도 VC는 최소한의 정보만을 제공할 수 있다. 하지만 이러한 VC는 시청자의 DID를 포함하고 있으며, 오로지 검증인인 시청조사단체만이 검증할 수 있기 때문에 외부에 공개될 경우 데이터의 완전성, 정확성, 일관성을 보장하는 무결성을 잃을 경우가 있다. 이로 인해 재구성된 시청 정보는 시청조사단체는 시청자의 VC를 검증한 ZKP 검증한 시청 정보 결과를 다시 자신의 DID로 치환한 VC_R 로 재구성하여 외부에 제공할 수 있도록 한다.

3.3 외부 공유를 위한 재구성된 시청 정보의 생성 방법

외부 공유를 위한 재구성된 시청 정보 자격 증명은 시청자의 DID를 시청조사단체의 DID로 치환되었기 때문에 어느 시청자인지 특정할 수 없는 인과 관계가 성립되지 않는 상태가 되며, 외부에 공유 되어도 정보의 무결성이 보장된 상태가 되며, 재구성된 시청 정보 생성 방법은 그림 4와 같다.

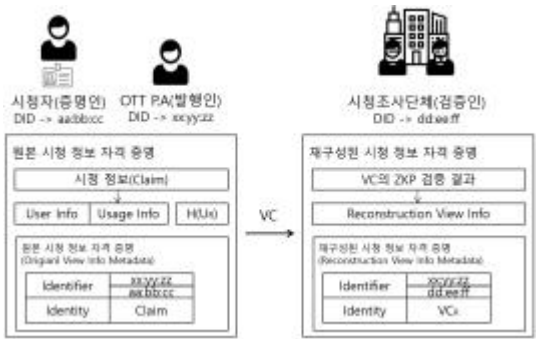

그림 4. 재구성된 시청 정보 생성 방법
Fig. 4. Reconstructed viewing information generation method

원본 시청 정보 자격 증명인 VC는 시청자의 시청 정보는 이용 정보 및 시청 정보로 생성한다. 이 중 이용 정보는 이름, 나이, 성별 등으로 시청률 산출에는 필요하지 않는 정보이다. OTT 플레이어 에이전트 환경에서 발생하는 시청 정보에서 시청 정보를 제외한 이용자 정보에서 $User\ Info$의 SHA-256 해시값인 $H(U_R)$를 추가적으로 생성한다. OTT 플레이어 에이전트는 원본 VC를 발행할 때 시청자의 $User\ Info$ 및 $H(U_R)$를 기반으로 생성하며, 검증 과정을 거친다. 그리고 시청조사단체는 ZKP 검증 결과를 바탕으로 VC_R를 생성한다.

3.3 재구성된 시청 정보의 외부 공유 방법

시청조사단체는 발행된 VC_R를 권리자 등 외부인에게 공유할 수 있으며, 그림 5와 같다.

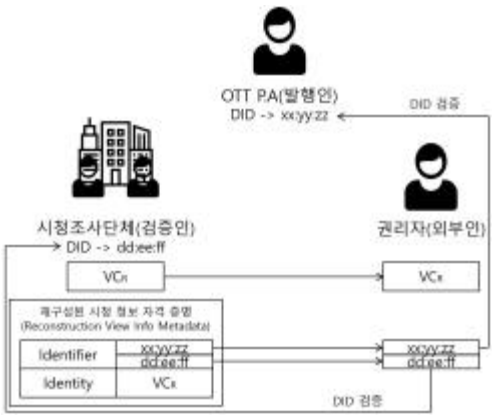

그림 5. 시청 정보의 외부 공유 방법
Fig. 5. External sharing method of viewing information

권리자 등은 제공받은 VC_R에 포함된 DID를 질의한다. DID는 두가지로 시청조사단체의 DID와 OTT 플레이어 에이전트의 DID로 두가지가 동일하면 VC_R의 유효성이 입증되어 시청 정보를 활용할 수 있다.

4. 영지식 증명 기반 OTT 플레이어 에이전트 환경의 시청 정보 생성 및 검증 방법

4.1 OTT 플레이어 에이전트 환경의 시청 정보 생성 및 검증 방법

시청조사단체가 시청 정보 수집을 위해서 시청자의 시청 정보를 VC로 수집한다. 이때 콘텐츠의 권리 행사를 위한 권리자 또는 권리 대행자는 본인 소유의 콘텐츠의 시청 분석을 위해서 오직 시청 정보만을 요구한다. 이때 시청 정보에 포함된 이용자 정보는 인물을 특정할 수 없도록 DID 등의 고유 식별 번호로만 구성한다. 이를 위한 영지식 증명 기반 OTT 플레이어 에이전트 환경의 시청 정보 생성 및 검증 전체 구성도는 그림 6과 같다.

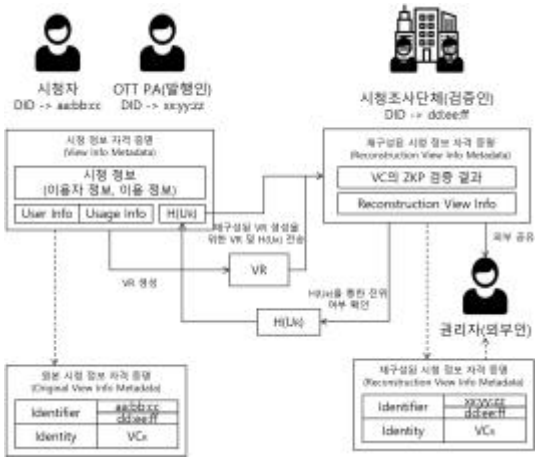

그림 6. OTT 플레이어 에이전트 환경의 이용정보 생성 및 검증 방법
Fig. 6. Method of generating and verifying usage information of OTT player agent environment

4.2 VC_R 생성 방법

먼저, OTT 플에이어 에이전트는 시청조사단체가 요청한 정보를 확인하는 것으로 시청조사를 위해 요구하는 자격 증명을 확인한다. 이때, 요청한 정보는 일정 기간 내 콘텐츠 별 정보 또는 모든 콘텐츠 정보 등으로 시청 조사를 어떻게 할 것인가에 따라 다르다. 해당 시청 정보를 VC로 생성한다. 이때 이용자 정보 및 이용 정보가 Claim이 되고 OTT 플레이어 에이전트가 발행하는 시청 정보가 VC가 된다. 이 과정에서 VC에 대해 $H(U_R)$을 생성한다. 생성된 $H(U_R)$을 시청조사단체에 제공한다. 시정조사단체는 모든 검증이 성공적으로 종료될 경우 시청 정보에 대한 영지식

증명 검증 결과를 획득한다. 또한, 권리자 등에 시청 정보를 제공할 수 있도록 재구성된 시청 정보 VC_R을 생성한다.

4.3 VC_R 검증 방법

시청조사단체는 권리자 등의 요청에 의해 적합한 VC_R을 제공한다. 시청조사단체 및 권리자 등 간에 VC_R의 영지식 증명을 수행한다. 권리자 등은 이 과정에서 사전에 영지식 증명 기반으로 검증이 끝난 시청 정보와 시청조사단체 및 OTT 플레이어 에이전트의 DID를 획득할 수 있으며, DID를 통해 시청조사단체와 OTT 플레이어 에이전트에 질의하여 VC_R의 유효성을 검증한다. 만일 시청 정보 및 DID가 상이하거나 시청 정보 발행 사실을 부정할 경우 VC_R은 유효하지 않는 것으로 판단하여, 시청조사 목적으로는 사용하지 않는다.

5. 결론

본 논문에서는 생성되는 시청 정보에 포함되어 있는 이용자 정보와 이용 정보에서 시청조사에 필요치 않는 이용자 정보를 영지식 기반으로 비식별화하고 외부 공유되었을 때 무결성을 검증할 수 있는 시청 정보의 재구성된 자격 증명을 이용한 영지식 증명 기반 OTT 플레이어 에이전트 환경의 시청 정보 비식별화 및 검증 방법을 연구하였다. 시청조사단체가 생성한 VC_R은 시청자의 개인정보 및 콘텐츠 이용정보 자격 증명에 대한 원본은 아니지만, 신뢰할 수 있는 영지식 증명으로 재구성되었기 때문에 외부에 공유할 수 있으며, 특히 권리자 및 권리 단체가 권리 행사를 충분히 할 수 있을 것이라고 사료된다.

추후 연구로는 실제 검증 과정을 통하여 VC_R에 대한 무결성을 보장하는 작업이 필요하다.

This work was supported by Ministry of Science and ICT and Institute for Information & communication Technology Planning & evaluation (IITP) (2022-0-00510)

참 고 문 헌

[1] D. Lee, “Comparative Study on Operational Optimization with Contactless e-Service Encounters”, International Business Education Review, Vol. 20, No. 5, pp.51-71, 2023. DOI : https://doi.org/10.38115/asgba.2023.20.5.5

[2] J. Lee, S. Cha, “A study on the survival strategy of K-content after entering the global OTT service platform domestic market”, Vol. 27, pp.149-173, 2022. DOI : 10.24174/jicc.2022.10.27.149

[3] Y. Kim, “OTT service usage status by generation”, KISDI STAT Report, Vol. 22, No. 07, pp.1-7, 2022, https://www.kisdi.re.kr/report/view.do?key=m2101113025790&masterId=4333447&arrMasterId=4333447&artId=657336

[4] G, Yoo, “KOCCA Global OTT Trend Report”, KOCCA, Vol. 2, pp.4-22, 2023. https://www.kocca.kr/globalOTT/vol02/html/main.html

[5] G. Min, “Audience rating survey outlook based on changes in viewing behavior”, Broadcast and Media Magazine, Vol. 27, No. 4, pp.35-39, 2022, http://www.kibme.org/resources/journal/20221123115024656.pdf

[6] Editorial Department, “Trends in standardization of IPTV viewing information measurement functions”, 2011, https://elec4.co.kr/m/article/articleView.asp?idx=1480

[7] S. Oh, J. Im, “Research on market response to integrated viewership ratings and utilization measures to revitalize the

terrestrial advertising market", 2015. https://www.kobaco.co.kr/site/adstat/board/report_etc_report/180?cp=6&sortOrder=BA_REGDATE&sortDirection=DESC&bcId=report_etc_report&baNotice=false&baCommSelec=false&baOpenDay=true&baUse=true
[8] M. Joo, "Personal information protection column-5] Domestic and international trends in personal information de-identification systems", Security News 2023. https://m.boannews.com/html/detail.html?mtype=2&tab_type=2&idx=119998
[9] Y. Min, "A Study on the Processing Method of pseudonym information considering the scope of data usage", Vol. 26, No. 5, pp.17-22, 2021. DOI : https://doi.org/10.9708/jksci.2021.26.05.017
[10] S. Kim, S. Jeon, "Big Data Integration Using Data De-identification", Journal of the Korean Intelligent Systems Society, Vol. 29, No. 3, pp. 235-241, 2019. DOI : 10.5391/JKIIS.2020.30.3.228
[11] S. Cha, T. Hsu, Y. Xiang, K. Yeh, "Privacy enhancing technologies in the internet of things: Perspectives and challenges", IEEE Internet of Things Journa, Vol.6, No.2, pp.2159-2187, 2019. https://ieeexplore.ieee.org/document/8515008

저 자 소 개

박병찬(Byeongchan Park)

2015.2 학점은행제 졸업
2018.2 숭실대학교 컴퓨터학과 석사
2023.8 숭실대학교 컴퓨터학과 박사
2023.9-현재 숭실대학교 초빙교수
<주관심분야> 저작권 보호 및 이용활성화

장세영(Seyoung Jang)

2018.2 평생교육원 학점은행 졸업
2021.6 숭실대학교 컴퓨터학과 석사
2023.2-현재 숭실대학교 컴퓨터학과 박사과정
<주관심분야> 저작권 보호 및 이용활성화

김석윤(Seok-Yoon Kim)

1980.2 서울대학교 전기공학과 졸업
1990.2 University of Texas at Austin Dept. of ECE 석사
1993.8 University of Texas at Austin Dept. of ECE 박사
1982-1987 ETRI 연구원
1993-1995 모토로라(Austin, Tx) 책임 연구원
1995-현재 숭실대학교 교수
<주관심분야> 시스템설계방법론, 저작권보호기술

김영모(Youngmo Kim)

2003.2 대전대학교 컴퓨터공학과 졸업
2005.2 대전대학교 컴퓨터공학과 석사
2011.2 대전대학교 컴퓨터공학과 박사
2012-현재 숭실대학교 교수
<주관심분야> 저작권 보호 및 이용활성화