가명화된 차량 운행기록의 가명결합 방안 연구

김은진[1], 강병훈[2]
[1]한국과학기술원 정보보호대학원 박사과정
[2]한국과학기술원 전산학부 교수
eunjin.kim@kaist.ac.kr, brentkang@kaist.ac.kr

A Study on Pseudonym Combination for Pseudonymized Vehicle Records

Eunjin Kim[1], Brent Byunghoon Kang[2]
[1]Graduate School of Information Security, KAIST
[2]School of Computing, KAIST

요 약

교통체계서비스는 안전한 교통 환경을 구축하는 것을 목표로 하여 차량, 도로, 기반 시설의 정보를 수집 및 처리하여 안전 교통정보를 제공한다. 교통체계서비스가 수집하는 차량 운행정보는 교통 안전 정보 외에도 다른 분야에서도 활용될 수 있으며 특히 다른 데이터와 결합하는 것으로 다양한 결과를 도출할 수 있어 연구, 통계 작성 등에 필요한 자료이다. 그러나 차량의 운행정보는 운전자의 개인정보를 포함하고 있어 운행정보 활용 시 가명화 및 가명결합이 필수적이다. 본 논문에서는 가명화된 운행정보를 가명결합하는데 발생하는 문제점을 설명하고 이러한 문제를 해결한 가명결합 방안을 연구하였다. 그 결과 교통체계서비스가 수집한 운행정보를 다른 기관의 데이터와 결합하여 활용할 수 있게 하여 개인정보를 보호하면서 데이터의 유용성을 활용하는데 기여할 것으로 예상한다.

1. 서론

최근 사물 인터넷과 5G 등 통신 기술의 발전으로 안전한 교통 환경 구축을 위한 교통체계서비스(C-ITS)[1] 구축 방안이 논의되고 있다. C-ITS는 차량과 시설물 등에서 실시간으로 정보를 수집하고 이를 통하여 다양한 교통정보를 공유하는 것을 목표로 한다. 수집된 정보는 실시간으로 AI 등의 기술로 처리되어 교통안전 정보를 생성하고 이를 통해 사고를 미리 방지할 수 있도록 한다.

그러나 이러한 시스템 모델은 차량의 운행기록을 수집하는 만큼 운전자의 프라이버시를 침해할 수 있다는 문제점이 있다. 차량 운행기록은 운행기록계나 운전자의 식별번호, 차량의 속도, 위치 정보 등 프라이버시를 침해할 수 있는 정보들을 포함하고 있다. 이러한 문제를 해결하기 위해 운행기록에서 운전자나 차량을 식별 가능한 정보를 제거하고 대신 가명을 활용하는 가명화 방안 연구가 존재한다[2][3][4]. 가명화는 운행기록에서 운전자 식별정보 등 프라이버시가 노출될 수 있는 정보를 제거하여 프라이버시 침해를 최소화할 수 있다. 따라서 교통체계서비스에서 운행정보의 수집 시 운행정보의 가명화가 필수적이다.

교통체계서비스에서 수집된 운행기록은 실시간 교통정보 분석뿐만 아니라, 교통안전 기술 연구, 정책개발, 수사자료 등 공익 목적으로 활용될 수 있다. 이때 운행기록을 다른 기관이 보유한 데이터와 결합하는 것으로 단일 정보로 얻을 수 없는 새로운 정보를 얻을 수 있다. 각각 다른 기관이 보유한 정보의 결합은 공통된 정보를 연결하여 데이터를 통합하는 것으로 수행할 수 있다. 그러나 데이터를 결합하는 과정에서 개인정보가 유출될 가능성이 있어 주의를 기울여야 한다. 예를 들어 운행기록정보와 보험정보를 결합할 경우, 차량번호 등 민감한 개인식별정보를 기준으로 하여 서로 다른 데이터를 결합할 수 있다. 이 과정에서 민감한 개인식별정보가 노출되어 프라이버시가 침해가 발생할 수 있다. 이를 해결하기 위하여 개인정보보호법에서는 서로 다른 기관이 소유한 데이터를 결합하여 사용할 때는 가명정보 결합을 사용할 것을 명시하고 있다[5][6]. 따라서 교통체계서비스가 수집한 운행기록을 가명결합 서비스를 활용하여 안전하게 활용하는 방안의 검토가 필요하다.

그러나 운행기록과 다른 기관의 데이터를 가명결합할 때 다음과 같은 문제점이 존재한다. 첫째, 운행정보는 교통체계서비스에 전송되기 전에 가명화 되어야 하므로 가명결합에 활용할 수 있는 정보가 거의 없어 어려움이 존재한다. 가명결합 절차는 결합

신청자와 결합키 관리기관과 협의를 통하여 실제 활용할 속성 이외의 동일 속성 데이터를 이용하여 결합키를 생성해야 한다. 이 동일 속성으로는 주로 개인정보를 담고 있는 식별정보가 포함된다. 그러나 이미 가명화된 운행정보에는 이러한 속성들이 제거되어 있다. 따라서 이미 가명화된 운행정보에서 결합키를 생성하는 방법 연구가 필요하다.

둘째, 운행정보의 가명은 가명 생성 전략에 따라 주기적으로 혹은 특정 조건에서 변경되어야 한다. 같은 가명을 지속해서 사용할 경우 공격자가 운행정보를 통해 운전자를 식별할 가능성이 있다. 공격자는 같은 가명을 사용하는 운행정보를 시간순으로 나열하는 것으로 운전자의 전체 혹은 부분 운행경로를 구할 수 있다[7][8]. 이때 만들어진 운행경로에 주거지, 회사 등 운전자를 추론할 수 있는 장소가 포함된 경우 운전자의 신원이 식별되고 프라이버시가 침해될 수 있다. 이를 피하기 위해서 운행정보의 주기적인 가명 변경이 필요하다. 따라서 주기적으로 혹은 특정 조건에서 갱신되는 가명을 이용하여 데이터를 결합 가능한 방법 또한 고려되어야 한다.

본 논문에서는 교통체계서비스가 수집한 운행기록을 안전하게 활용하기 위하여 운행기록을 위한 가명결합 방안에 관하여 연구한다. 본 논문에서는 먼저 운행정보에 관련된 가명화 연구에 대해서 살펴보본다. 또 가명화된 운행정보의 가명결합 방안에 논의하기 전에 이 과정의 이해를 돕기 위하여 가명정보결합에 대해 설명한다. 그 이후 가명정보결합을 활용할 수 있도록 운행정보의 가명을 추후에 결합키 생성에 활용하는 가명 생성 방법을 설계한다. 마지막으로 제안한 가명화 방법이 적용된 운행정보의 가명결합 절차에 대해서 논의한다.

2. 관련 연구

운전자의 프라이버시를 보호하기 위하여 다양한 차량운행정보 공유 환경에서 가명을 지속적으로 변경하는 방안들이 연구되고 있다. 운행정보를 공유하는 VANET, V2X 등의 차량간 네트워크에서는 주기적인 차량 운행정보(안전 메시지, 비콘 등)를 브로드캐스팅하여 주변 차량 등에 전송한다. 이러한 운행정보에는 ITS에서 수집하는 운행정보와 마찬가지로 차량의 위치 정보 등 운전자의 프라이버시를 침해할 수 있는 정보가 포함되어 있다. 따라서 운행정보 공유를 통한 프라이버시 침해를 막고자 운행정보의 식별자를 제거하고 가명을 사용하는 연구가 진행되고 있다.

특히 가명을 안전하고 효과적으로 변경하기 위하여 다양한 연구가 진행되어오고 있다. 믹스존은 특정 영역에 들어가고 나가는 차량들을 혼동시키도록 영역 내의 차량의 가명을 변경하는 방법이다[1][3][9]. 이러한 방법을 통하여 공격자가 운전자의 다음 위치를 추론하지 못하는 것을 목표로 한다. 또 다른 가명 변경 방안으로 다수의 차량이 정차하거나 방문하는 장소에서 가명을 변경하는 방안이 제시되었다[1]. 그 외에 차량이 일정 이하의 속도로 주행할 때 운행정보의 전송을 멈추고 이때 가명을 변경하는 방법[3], 랜덤하게 생성된 가명이 아닌 공개키 암호 방식 등의 암호 기술으로 생성된 가명을 이용한 방법[8][9] 등이 제안되었다.

3. 가명정보결합

가명정보 결합이란 각각의 개인정보 취급자가 보유한 개인정보를 가명 처리하여 결합하는 것을 의미한다[6]. 개인정보보호법은 통계, 연구 등으로 개인정보를 활용할 때 개인정보 보호를 위하여 가명 처리하여 활용하도록 하고 있다. 가명정보결합은 개인정보가 포함된 데이터를 결합하여 활용할 때 개인정보 침해 위험을 경감시키면서 단일 데이터에서 얻지 못한 유용한 정보를 도출할 수 있게 한다.

안전한 가명정보 결합을 위하여 데이터의 결합신청자는 결합전문기관과 결합키관리기관을 통하여 데이터를 결합하여야 한다. 먼저 데이터 결합을 협의한 상호 기관들은 결합키기관과 협의를 통하여 결합키 생성정보를 선택한다. 결합신청 기관은 결합키를 생성하기 위하여 결합키 생성정보를 선택하고 각 데이터 항목마다 결합키기관이 생성한 솔트 값과 함께 해시값을 생성한다. 결합키 생성정보는 이름, 전화번호, 주소 등의 개인정보 등 각 데이터 항목을 식별할 수 있는 속성이 선택된다. 결합키가 생성된 후에 결합신청 기관은 각 데이터 항목의 결합키와 일련번호를 결합키관리기관에 전송한다. 결합키관리기관은 각 데이터의 결합키와 일련번호를 이용하여 결합키 연계정보를 생성한다. 결합키 연계정보는 데이터 항목의 일련번호를 서로 매핑하여 상대 기관의 연관 데이터 항목과 연결한 정보이다. 결합전문기관은 생성된 결합키 연계정보를 이용하여 결합신청기관으로부터 받은 결합 대상 정보를 결합한다. 그 후 추가 가명처리 및 반출심사를 거쳐 결합된 데이터를 결합신청 기관으로 반출한다. 이때 결합전문기관은 관계

중앙행정기관의 장이 지정하는 전문기관이어야 한다. 이러한 과정을 통하여 정보 주체가 식별되지 않고 가명화된 상태에서 안전하게 데이터를 활용할 수 있다.

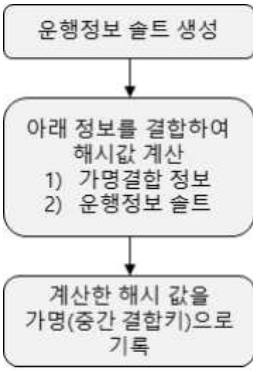
4. 가명결합 가능한 운행정보 가명화

본 논문에서는 가명화된 운행정보를 가명결합하기 위해서 운행정보의 가명을 중간 결합키로 활용하는 방법을 제안한다. 운행정보는 교통체계서비스에 전달되기 전에 가명화 된다. 이 때문에 교통체계서비스가 보관하고 있는 운행정보에는 결합키 생성을 위한 식별자가 존재하지 않는다. 이러한 문제를 해결하기 위해 본 논문에서는 결합키 생성에 활용될 것으로 예상되는 속성으로 가명을 생성하는 방법을 제안한다. 이러한 가명은 결합키 생성을 위한 데이터 속성은 아니나, 이를 대신하여 결합키 생성에 활용할 수 있도록 처리된 값이다. 따라서 이를 중간 결합키라고 부르기로 한다. 또한 가명을 이용한 운행정보 연결을 막기 위하여 생성된 중간 결합키가 갱신될 수 있도록 설계한다.

가. 중간 결합키 생성

가명화된 운행정보의 가명결합키를 생성하기 위하여 중간 결합키를 생성한다. 가명결합키를 생성하기 위해서 가명결합을 희망하는 기관들은 결합키 생성 정보를 합의하여야 하고 이런 정보는 주로 차량번호, 운행기록계 시리얼 번호 등 개인식별정보가 사용된다. 그러나 교통체계서비스가 저장하고 있는 운행기록은 이미 가명화가 되어 개인식별정보가 제거되어 있어 결합키를 생성하는데 어려움이 있다. 이를 해결하기 위하여 운행정보의 중간 결합키를 최소한 구성의 개인식별정보로하여 해시값을 구하는 것으로 생성한다. 그림 1은 운행정보의 가명, 즉 중간 결합키를 생성하는 방법을 나타낸다. 미리 지정된 가명결합 정보를 이용하여 해시값을 계산하여 가명, 즉 중간 결합키를 생성한다. 생성된 중간 결합키는 운행정보의 가명키를 대체하여 가명 역할 또한 수행한다.

먼저 가명결합 시에는 운행정보에 저장된 중간 결합키를 활용하여 결합키를 생성한다. 해시는 해시값의 이전 메시지를 모르는 상태에서 이전 메시지에 끝에 새로운 메시지를 추가하는 것이 가능한 특성이 있다. 이러한 해시의 특성을 이용하여 해시로 생성된 중간 결합키에 관리기관이 선택한 솔트를 추가하여 결합키를 생성한다. 그 결과 상대 결합신청기관은 대상 운행정보의 결합키와 같은 결합키를 개인정보 노출 없이 생성할 수 있다.



(그림 1) 운행정보 가명(중간 결합키) 생성 방법

나. 중간 결합키 갱신

가명결합 가능한 운행정보는 운행정보 연결을 통한 운전자를 식별을 막기 위하여 중간 결합키가 주기적으로 변경될 필요성이 있다. 이를 위하여 중간 결합키는 갱신 가능해야 한다. 본 논문에서는 중간 결합키를 갱신하기 위하여 운행정보의 솔트 값을 생성하고 이를 이용하여 중간 결합키를 생성한다. 운행정보는 그림 1과 같이 가명 갱신을 위하여 솔트를 생성하고 이 값을 가명결합 정보와 함께 중간 결합키 생성 시에 활용한다. 해시 값은 원본 메시지가 조금이라도 변경되면 그 값이 변화하는 특징이 있다. 따라서 중간 결합키는 운행정보의 솔트 값이 변화할 때 변화하게 되어 중간 결합키가 갱신 된다. 이 솔트 값은 랜덤 등의 방법을 활용하여 생성하여 주기적으로 혹은 특정 조건에서 갱신하도록 한다. 생성된 솔트 값은 운행정보에 하나의 속성으로 포함하여 추후 솔트가 사용된 기간과 함께 상대 결합신청 기관에 제공한다. 상대 결합신청 기관은 운행정보의 솔트 값을 받아 결합키 생성에 활용하는 것으로 대상 운행정보와 동일한 결합키를 생성할 수 있다. 이러한 특징을 이용하여 동일한 가명의 운행정보 연결을 통한 운행정보의 개인식별 위험을 줄일 수 있다.

5. 운행정보의 가명결합 절차 고려사항

본 논문에서 제안한 가명결합 가능한 운행정보 가명화 방법을 실제로 적용할 때 가명키 생성 시 제한사항이 발생한다. 가명결합을 위해서는 중간 결합키를 이용하여 생성된 결합키와 결합할 데이터의 결합키는 동일해야한다. 5장에서는 이러한 점을 고려하여 가명화된 운행정보의 가명결합 절차에 대해서 논의한다.

먼저 가명결합 시 결합키 생성정보를 협의할 때 기생성된 중간 결합키를 고려해야 한다. 결합신청기관은 결합키 생성정보를 합의할 때 중간 결합키(운행정보의 가명) 생성 시 사용된 속성 정보를 필수적으로 선택해야 한다. 운행정보의 중간 결합키는 가명화 이전에 선택된 정보를 기반으로 생성되어 이에 대한 해시값을 가지고 있다. 가명결합 시에는 이미 생성된 중간 결합키의 생성정보 구성을 변경할 수 없으며 결합키관리기관이 생성한 솔트나 추가적인 속성 정보의 추가만 허용된다. 따라서 상대 결합신청기관은 운행정보의 중간 결합키 생성 시 사용된 속성들을 같은 순서로 활용하여 운행정보와 같은 결합키를 생성할 수 있도록 해야 한다.

또한 가명결합 시 중간 결합키 갱신을 위하여 사용된 운행정보의 솔트 정보는 상대 기관에 공유되어 결합키 생성시 활용되어야 한다. 운행정보의 솔트 정보는 중간 결합키와 함께 운행정보에 포함되고 중간 결합키 생성 시 해시 계산에 활용된다. 따라서 운행정보를 공유하는 교통체계서비스는 운행정보 솔트 정보와 각 솔트의 사용 기간을 상대 기관에 공유하여 동일 결합키를 생성할 수 있도록 해야 한다.

6. 결론

본 논문에서는 교통체계서비스에서 이미 가명화된 운행정보를 위한 가명결합 방안을 논의하였다. 결합키 생성 시에 주로 사용되는 운행정보의 개인식별 속성을 활용하여 중간 결합키를 생성하고, 이를 운행정보의 가명으로 활용하는 방법을 제안하였다. 또한 가명을 이용하여 운행경로 복구 및 운전자 식별 시도를 막기 위하여 갱신 가능한 중간 결합키를 제안하였다. 마지막으로 제안한 중간 결합키를 활용하여 가명결합을 수행할 때 고려해야 할 사항을 토의하였다. 가명결합 절차는 데이터 주체의 개인정보를 보호하면서 연구, 통계 작성 등의 공익 목적에서 데이터를 유용하게 활용할 수 있도록 지원한다. 본 논문은 프라이버시 보호된 데이터의 활용을 위하여 교통체계서비스의 운행정보를 위한 가명결합 방안을 제안하였다. 이를 활용하여 운전자의 개인정보를 보호하는 동시에 도로 교통 분야의 연구를 지원하는 것으로 안전한 교통 환경을 구축하는데 기여할 수 있을 것이다.

* 본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 22CTAP-C163794-02).

참고문헌

[1] 방송통신위원회, 한국인터넷진흥원, “차세대 지능형 교통시스템 (C-ITS) 기술동향”, 위치정보사업 동향 보고서 4월, 2023, https://www.kisa.or.kr/post/fileDownload?menuSeq=20204&postSeq=189&attachSeq=1&lang_type=KO (Last accessed:2024.04.24.)

[2] R. Lu, X. Lin, T. H. Luan, X. Liang and X. Shen, "Pseudonym Changing at Social Spots: An Effective Strategy for Location Privacy in VANETs," in IEEE Transactions on Vehicular Technology, vol. 61, no. 1, pp. 86-96, Jan. 2012

[3] L. Buttyán, T. Holczer, A. Weimerskirch and W. Whyte, "SLOW: A Practical pseudonym changing scheme for location privacy in VANETs," 2009 IEEE Vehicular Networking Conference (VNC), Tokyo, Japan, pp. 1-8, 2009

[4] J. Benin, M. Nowatkowski and H. Owen, "Unified pseudonym distribution in VANETs," 2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications, Niagara Falls, ON, Canada, pp. 529-533, 2010

[5] 개인정보보호법, https://law.go.kr/법령/개인정보보호법 (Last Accessed: 2024.04.24.)

[6] 가명정보결합 종합지원시스템. https://link.privacy.go.kr/nadac/organ/introData.do (Last Accessed: 2024.04.24.)

[7] Abdelwahab Boualouache, Sidi-Mohammed Senouci, Samira Moussaoui, “VLPZ: The Vehicular Location Privacy Zone”, Procedia Computer Science, Volume 83, Pages 369-376, 2016

[8] Siham Bouchelaghem, Mawloud Omar, “Secure and efficient pseudonymization for privacy-preserving vehicular communications in smart cities”, Computers & Electrical Engineering, Volume 82, 2020

[9] J. Freudiger, M. Raya, M. Felegyhazi, P. Papadimitratos, and J.-P. Hubaux, “Mix-zones for location privacy in vehicular networks,” in Proceedings of the 1st International Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS 07), 2007