위치기반 서비스에서 프라이버시를 위한 연속질의와 쿼리 로그 익명화 기법

박소미, 배주호 ○, 박 석
서강대학교 컴퓨터공학과
thathal@sogang.ac.kr juho@sogang.ac.kr spark@sogang.ac.kr

Location Anonymization Techniques of Continuous Query and Query Log for Privacy in Location-Based Services

Somi Park, Juho Bai ○, Seog Park
Dept. of Computer Science & Engineering, Sogang University

위치기반 서비스 이용의 확산은 GPS가 탑재된 모바일 기기들의 발전으로 인해 가속화되고 있다. 위치기반 서비스란 이동 중인 사용자의 위치정보를 관련된 타 정보와 결합하여 사용자가 원하는 부가적인 응용서비스를 제공하는 서비스이다. 위치기반 서비스 제공자가 사용자의 정확한 위치를 항상 파악한다면 서비스 품질은 높아지지만, 서버가 공격을 당하거나 정보들이 누출이 된다면 사용자들의 위치 정보 역시 누출될 가능성이 있다.

따라서 사용자의 위치 정보 누출을 막기 위해 Location k-anonymity가 제안되었다.[1] 이 기법은 주변에 있는 k-1명 이상의 사용자들이 포함되는 영역을 서버로 보냄으로써 서버를 공격한 공격자가 질의자의 정확한 위치를 알아낼 수 있는 확률은 1/k 이하로 낮추어 준다. P. Kalnis et al.[2]에서는 Hilbert Curve Order를 사용하여 사용자들의 위치를 정렬, k명 이상의 그룹으로 나누고, 그룹에 속한 사용자들을 포함하는 최소 경계 사각형을 ASR로 정한다. 이 기법은 생성되는 영역의 크기에 대한 제한이 없고 사용자들의 움직임에 따라 리스트를 재정렬해야 한다. M. F. Mokbel, C. Chow와 W. G. Aref의 연구[3]에서는 그리드 기반의 피라미드 구조를 사용하는 기법을 제안하였다. 수평 혹은 수직의 이웃셀과 해당 셀에 포함된 사용자가 k명 이상인 경우 해당되는 이웃셀과 기존의 셀이 포함된 영역이 ASR이 되며, 만약 이웃셀을 포함하더라도 k명을 만족하지 못하는 경우, 피라미드 구조의 상위 단계에서 동일한 프로세스를 재귀적으로 실행하여 k명을 만족하는 영역을 탐색하게 된다. B. Bamba와 L. Liu의 연구[4]는 [3]와 비슷하게 그리드 기반의 기법을 사용하는데 사용자가 포함된 셀의 열 방향으로 증가시키며 k명 이상을 포함하는지 살펴본다. 만약 만족하지 못할 경우 행 방향으로 셀을 증가시키며 만족하는 영역을 탐색한다. 이 기법은 k-anonymity뿐만 아니라 사용자 주변의 POI에 대한 l-diversity를 고려한다. 따라서 ASR을 생성할 때, 두 가지를 동시에 만족하는 영역을 고려하게 되므로, 영역의 크기가 다른 기법들에 비해 커질 수 있다. 또한 [4]과 [3]은 충분히 많은 사용자들이 서비스에 등록되어 있어야 하며, 사용자들의 위치가 고르게 분포되지 않은 경우 적정 수준의 영역을 생성하는데 실패할 수 있다. 이러한 중앙집중 방식은 중앙 서버, 즉 위치기반서비스의 익명화 서버를 절대적으로 신뢰할 수 있다는 전제 하에 가능하다. 그러나 이 모델의 취약점은 중앙집중 익명화 서버에 과부하 및 병목현상이 생길 수 있다는 것이다. 따라서 기존의 연구는 k명의 사용자가 포함된 영역(Anonymized Spatial Region; ASR)을 효율적이고 안전하게 생성하는 것에 초점이 맞추어져 있다.

그러나 기존의 연구들은 사용자의 이동성과 연속적인 질의에 대한 고려가 부족하다. 사용자가 이동 중에 연속하여 질의를 요청한다면 이때 생성되는 익명화 영역들에는 항상 사용자가 포함된다. 만약 일련의 익명화 영역들을 교차할 수 있다면 공격자는 실제 사용자에 대한 정보를 얻을 수 있다. 또한 한 구성요소를 전적으로 믿을 수 있다는 가정은 실제적으로 불가능하며 특히, 익명화 서버에 저장되는 쿼리 로그는 실제 사용자의 위치 정보가 내포될 수 있다.

또한, 이러한 쿼리 로그에 대한 프라이버시는 내부자에 의한 노출 및 오·남용될 가능성이 있다. 이와 같은 상황을 고려하여 본 논문은 기존의 기법들의 문제점인 연속적인 질의에 대해 교차그룹

공격을 방지할 수 있는 새로운 익명화 기법을 제안하고, 익명화 서버에 저장되는 쿼리 로그의 프라이버시 문제를 인지하여 안전하고 효율적으로 저장할 수 있는 기법을 적용하여 먼저 기존 기법의 문제점인 연속적인 질의에 대한 교차공격을 방지할 수 있는 새로운 익명화 기법을 제시한다.

교차그룹 공격은 익명화된 두 개 이상의 그룹에 대해 교집합에 속한 원소의 개수가 k보다 낮아질 경우 발생한다. 따라서 연속되는 질의에 대해서 사용자가 동일한 익명화 그룹을 생성할 수 있다면 이러한 교차 그룹 공격을 방지할 수 있다. 본 논문에서 제안하는 Entourage Generation 기법은 한 사용자가 일정한 시간(Threshold t) 안에 연속하여 질의를 요청한다면 이전의 질의와 동일한 익명화 그룹을 이용하여 ASR을 생성한다. 이 기법을 통하여 익명화 프로세스는 연속적인 질의에 대한 익명화 그룹이 동일하게 생성될 수 있도록 하여 그 그룹들의 교집합이 Location k-anonymity를 만족할 수 있도록 한다.

그 다음으로는 익명화 서버의 쿼리 로그에 대해 사용자의 질의에 대한 프라이버시를 동시에 고려하는 기법을 제안한다. 서버에 저장된 기존의 데이터를 보존하면서 사용자의 실제 위치를 공격자가 알 수 없게 하기 위하여 익명화 그룹의 컬럼별로 데이터의 위치를 재배열 한다. 데이터의 재배열 후 실제 질의자의 데이터가 위치한 부분을 암호화하여 키 값으로 저장한다. 차후에 원본 데이터가 필요한 경우는 이 키 값을 사용하여 실제 질의자의 데이터를 복구할 수 있다. 원본 데이터 위치에 대한 키 값은 One-way Trapdoor Function을 사용하고, 생성되는 키를 이용하여 섞여진 테이블에서 실제 데이터의 위치를 파악할 수 있고 원본 데이터로 복구가 가능하다. 이 기법을 통하여 공격자가 서버에 저장된 쿼리 로그에 접근 가능한 경우, 어떠한 사용자가 실제 질의자이고, 어디에서 무엇을 찾고자 하는지를 모호하게 만들어 프라이버시를 보호한다.

본 연구는 기존 기법들의 문제점을 해결하기 위한 익명화 기법을 이용하여 수행시간을 약 50% 단축하고, 사용자 수나 익명화에 필요한 값이 증가 하더라도 익명화 그룹 범위를 일정 범위로 제한하는 방법을 제시한다

참고 문헌

[1] M. Gruteser and D. Grunwald, “Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking,” Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, pp. 31-42, 2003.

[2] P. Kalnis, G. Ghinita, K. Mouratidis et al., “Preventing Location-Based Identity Inference in Anonymous Spatial Queries,” IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 12, pp. 1719-1733, 2007.

[3] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, “The New Casper: Query Processing for Location Services Without Compromising Privacy,” Proceedings of the 32nd International Conference on Very Large Data Bases, pp. 763-774, 2006.

[4] B. Gedik and L. Liu, “Location Privacy in Mobile Systems: A Personalized Anonymization Model,” Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, pp. 620-629, 2005.