

실내공간에서 계층 구조를 이용한 K-익명화

K-Anonymity using Hierarchical Structure in Indoor Space

김 준 석* 이 기 준**
Joon Seok Kim Ki Joune Li

요 약 실내공간이 점차 복잡해짐에 따라 실외뿐만 아니라 실내에서도 LBS (Location Based Service)의 수요가 증가하고 있다. 그러나 실생활에 편의를 주는 LBS의 이면에는 개인의 위치 노출과 이로 인한 프라이버시 침해의 문제가 포함되어 있다. 위치 K-익명화 기술은 적어도 K-1명의 다른 사용자를 포함시키는 ASR (Anonymized Spatial Region)을 생성하여 위치를 은폐하여 이러한 프라이버시 문제를 해결하는 대표적인 연구 분야이다. 하지만 기존 연구들은 대부분 유클리디안 거리를 기반으로 하는 실외 공간에서 이루어 졌으며, 구조물에 의해 제약이 있는 실내공간에 적용시키기에는 한계점이 존재한다. 이에 본 논문에서는 실내공간의 구조와 위치 표현 방법을 고려하여 위치 은폐를 위한 K-익명화 방법을 제안한다. 위치의 은폐를 위해 실내공간의 계층 구조를 생성하여 이 구조에서 K-1명의 다른 사용자를 포함하도록 노드를 확장시키는 방법을 소개한다. 또한 제안된 방법의 성능을 분석하기 위하여 K와 계층구조의 특성에 따른 비용모델도 함께 제시한다.

키워드 : 실내공간, K-익명화, 계층구조

Abstract Due to complexity of indoor space, the demand of Location Based Services (LBS) in indoor space is increasing as well as outdoor. However, it includes privacy problems of exposing personal location. Location K-anonymity technology is a method to solve the privacy problems with cloaking their locations by Anonymized Spatial Region(ASR). It guarantees K users within a region containing the location of a given user. However previous researches have dealt the problems based on Euclidean distance in outdoor space, and cannot be applied in indoor space where there are constraints of movement such as walls. For this reason, we propose in this paper a K-anonymity for cloaking indoor location in consideration of structures and representation of indoor space. The basic concept of our approach is to introduce a hierarchical structure as ASR for including K-1 users for cloaking their locations. We also proposed a cost model by K and attributes of hierarchical structure to analyze the performance of the method.

Keyword : Indoor Space, K-Anonymity, Hierarchical Structure

1. 서 론

최근 측위 기술의 발달과 스마트 폰의 보급으로 위치기반 서비스(Location Based Service; LBS) 산업이 발달하고 컨벤션 센터나 백화점, 지하철역과 같이 복잡한 실내 구조를 가진 건물의 수가 증가함에 따라 실외뿐만 아니라 실내 내비게이션과 같은 실내 공간에 대한 LBS의 수요가 늘어나고 있다. 사용자들은 모바일 단말기를 통해 “A 브랜드 매장으로

가는 최단 경로를 알려 달라”와 같이 질의하여 실외와는 독립적으로 실내 위치기반 서비스를 받을 수 있다.

그러나 이러한 서비스들을 이용하게 되면 사용자의 개인의 위치 정보가 노출될 가능성이 있다. 서버-클라이언트 구조를 가진 LBS의 경우, 사용자가 서비스를 이용하기 위해 질의를 서버에 요청할 때 사용자가 원치 않게 자신의 위치를 서버로 보내야 하기 때문이다. 예를 들어, 백화점에서 “A 브랜드

† 본 연구는 두뇌한국21사업과 국토해양부 첨단도시개발 연구개발사업의 연구비지원(11첨단도시G11)에 의해 수행되었습니다.

* 부산대학교 컴퓨터공학과 박사수료 joonseok@pnu.edu

** 부산대학교 컴퓨터공학과 교수 lik@pnu.edu(교신저자)

매장 경로를 알려 달라”와 같은 질의를 하였을 때, 누군가가 A 브랜드 매장에서 이 시간의 결제 고객 명단을 함께 가지고 있다면, 의도치 않게 제 3자에게 사용자를 유추할 수 있는 정보를 노출하게 된다. 따라서 실외뿐만 아니라 실내에서 안전한 위치기반 서비스 사용을 위한 개인 정보 보호 방법이 필요하다.

위치정보를 공개하거나 위치기반 질의를 할 때, 개인의 프라이버시를 보호하기 위한 많은 연구가 활발히 이루어져 왔다[2, 3, 4, 7, 10, 15, 16]. 기존의 연구들 중에는 사용자의 위치를 확장시켜 은폐시키는 위치 K -익명성(Location K -anonymity) [4]에 관한 연구들이 존재한다. 위치기반 질의를 요청할 때, 사용자의 정확한 위치 정보 대신에 위치를 확장시켜 주변에 있는 최소 $K-1$ 명 다른 사용자들이 포함되는 익명화 공간영역(Anonymized Spatial Region: ASR)을 보내면서 위치를 은폐하는 것이다. 그런데 기존의 위치를 은폐하는 방법은 실외 환경에 적합한 유클리디안 거리를 기반으로 K -익명화를 만족하도록 영역을 확장하는 것으로 실내 환경에 대한 고려가 되어 있지 않다. 벽이나 천장, 바닥과 같은 구조물에 의해 이동의 제약을 받는 실내공간에서는 위치와 거리에 대한 특성이 다르므로 기존의 방법을 그대로 적용하기에는 한계가 있다.

본 논문에서는 실내공간에서 K -익명화로 사용자의 위치를 은폐시키는 방법을 제안한다. 실내공간의 경우 (x, y, z) 와 같은 3차원 좌표보다 ‘203호’, ‘2층 복도’와 같이 사람이 인지하기 쉬운 기호로 위치를 표현한다[8, 14]. 실내공간의 계층 구조가 존재할 경우, 사용자의 위치를 나타내는 방과 같은 공간을 하나의 노드로 하여, 이 노드와 다른 노드들을 포함하는 상위 노드를 선택함으로써 위치를 확장할 수 있다. 우리는 적어도 $K-1$ 명의 다른 사용자를 포함하도록 계층 구조에서 최소 레벨의 노드를 선택함으로써 효과적으로 은폐시키는 방법을 소개한다.

본 논문은 다음과 같이 구성된다. 2장에서는 K -익명성의 위치 은폐 기법과 실내공간을 계층 구조화한 기존 연구를 소개하며 연구 배경을 서술한다. 3장에서는 실내공간에서 위치를 표현하는 방법과 계층 구조를 나타내기 위한 계층 그래프를 정의한다. 4장에서는 위치 K -익명성을 위한 은폐를 계층 그래프에 적용한 방법과 비용 모델을 제시한다. 마지막으로 5장에서는 본 논문을 요약하고 향후 연구에 대해 논의한다.

2. 관련 연구 및 연구 동기

2.1 K -익명성

사용자의 위치 정보를 보호하기 위하여 연구된 은폐 기법들은 데이터 중심의 은폐 방법과 공간 중심의 은폐 방법으로 구분할 수 있다[3].

Clique 은폐 기법[2]은 영역을 기반으로 하는 공간 중심의 은폐 방법이다. 사용자가 서비스를 받기 위해 질의를 서버로 보낼 때, 사용자에게 해당하는 노드가 그래프에 생성된다. 그리고 사용자의 위치는 x 축, y 축에 평행하게 Δx , Δy 만큼 확장시켜 사각형 (constraint box)으로 확장하면서 자신의 사각형 안에 질의를 요청한 다른 사람의 위치가 포함되는지를 확인한다. 만약 다른 사람의 위치가 존재하고, 상대방의 사각형 안에 자신의 위치 또한 포함된다면 그래프에서 두 사용자에게 해당하는 노드는 에지로 연결된다. 그래프에서 에지로 연결된 노드가 K 개 이상이 되면 이를 K -clique라 하고, 이 K 개의 위치를 포함하는 ASR가 생성되어 질의를 수행한 후 결과를 각각의 요청자에게 전달한다. 이 때, 질의가 K -clique가 될 때까지 무한정 기다릴 수 없기 때문에 최대 시간 동안 K -clique를 찾지 못할 경우 해당 요청은 서비스 되지 않는다. Clique 은폐 기법은 응답 지연이 발생하거나 서비스를 받지 못하는 단점이 존재한다.

중심 은폐 기법[7]은 거리를 기반의 가장 기본적인 은폐 방법이다. 질의 요청자를 중심으로 가장 가까운 위치에 있는 $K-1$ 명을 찾아, 모두 포함할 수 있는 ASR을 생성한다. 하지만 중심 은폐 기법의 경우, “center-of-ASR” 공격에 취약하다. 요청자의 위치를 중심으로 ASR을 생성하기 때문에 요청자가 ASR의 중심에 위치할 확률이 매우 높다. 그래서 이러한 취약점을 보완하기 위해 중심 은폐 기법을 변형시킨 NN-은폐 기법(Nearest Neighbor Cloak) [7]이 존재한다. NN-은폐 기법에서 ASR은 다음과 같이 만들 수 있다. 우선 질의가 들어왔을 때 질의를 요청한 요청자를 중심으로 가장 가까운 $K-1$ 명의 이웃을 찾는다. 그리고 요청자를 포함한 K 명 중에 한 명을 랜덤하게 선택하고 선택된 한 명을 중심으로 가장 가까운 $K-1$ 명을 찾은 후, 랜덤하게 선택된 한 명을 중심으로 ASR을 만든다. 이 때 ASR에 질의를 요청한 요청자가 포함되지 않을 경우 요

청자까지 포함하도록 ASR을 확장시켜 준다.

Casper[10]는 가장 대표적인 격자 기반의 위치 K-익명성 방법이다. Casper는 익명화가 적용될 전체 공간을 격자로 나눈 후에 쿼드-트리(Quad-Tree) [12]와 유사하게 계층을 가지는 피라미드의 자료구조로 구성한다. 피라미드의 가장 높은 레벨은 전체 공간을 나타내고, 피라미드의 가장 낮은 레벨은 전체공간을 가장 세밀하게 나눈 공간을 나타낸다. 사용자가 질의를 서버로 요청하면 Casper는 가장 낮은 레벨에서 요청자가 위치한 셀을 탐색한다. 해당 셀에 K명 이상의 다른 사용자가 존재한다면 해당 셀이 ASR이 된다. 반면, 요청자가 위치한 셀이 K-1명의 다른 사용자들을 포함하지 않을 경우 Casper는 그 셀로부터 수평 혹은 수직의 방향에 위치해 있는 이웃 셀에 존재하는 사용자들을 찾아 요청자가 있는 셀의 사용자들의 수와 합하여 K명을 만족하는지 확인한다. 만약 만족한다면 찾아낸 이웃 셀들의 조합이 ASR이 되며 만족하지 못할 경우 Casper는 피라미드 구조에서 요청자가 존재하는 셀의 상위 레벨의 셀을 선택한다. 그리고 K명을 만족할 때 까지 위의 과정을 반복한다.

구간 은폐 기법[4]은 Casper와 유사하게 전체 공간을 같은 크기의 사각형으로 나누어 쿼드-트리와 같은 구조로 만든다. 하지만 구간 은폐 기법은 ASR을 계산할 때 같은 레벨의 이웃 셀을 고려하지 않고 바로 현재 위치 셀의 상위 레벨의 셀을 선택한다. Casper와 구간 은폐 기법을 비교했을 때, 확실히 Casper가 더 최소한의 넓이를 가지는 ASR을 생성한다. 하지만 [8]에서 증명되었듯이 구간 은폐 기법이나 Casper는 데이터가 전체 공간에 균일하게 분배되어 있어야 효과적인 방법이다.

위와 같은 단점을 보완하기 위해 힐버트 커브 정렬(Hilbert Curve Order)을 사용하는 힐버트 은폐 기법[10]이 존재한다. 2차원으로 나타낸 사용자들의 위치를 힐버트 커브를 이용하여 1차원으로 정렬한 후, K명씩 버킷을 만든다. 그리고 질의를 요청한 사용자가 포함된 버킷을 선택하고, 버킷에 속한 K명의 위치를 모두 포함하는 ASR을 만든다.

2.2 실내 공간의 계층 구조

이 절에서는 실내공간을 계층 구조화 한 기존의

연구를 살펴본다. 다양한 어플리케이션에서 실내공간을 여러 단계의 계층 구조로 나타내는 연구가 진행되었다. 특히 내비게이션과 관련된 연구에서 단순한 모델에서 점점 세밀한 모델로 나타나도록 방법이 연구되고 있는데, 이를 공간의 계층 구조로 볼 수 있다[5, 11, 13]. 이러한 연구 중에서 형식적인 방법으로 실내공간을 계층 구조로 나타낸 연구는 [13]이다. [13]은 실내공간의 평면도를 기초로 하여 공간과 공간의 연결정보를 가지고 기본 그래프를 만든 후 이 기본 그래프를 변형시켜 계층을 만드는 방법을 제안한다. 이 방법은 실내에서 이동객체가 이동할 때 방과 방으로 바로 연결되는 것이 아니라 방으로 들어가기 전에 주요 복도나 홀을 거쳐서 가야한다는 특성을 이용한다.

[13]에서 제안하는 방법론을 살펴보면, 실내공간의 평면도를 이용하여 벽과 같은 경계로 둘러싸인 방을 노드로 나타내며, 연결성이 있는 공간끼리 에지로 연결하여 연결정보를 기초로 한 기본 그래프를 만든다. 만들어진 기본 그래프에서 색칠된 노드를 선택하는데 색칠된 노드는 (1) 현관과 같이 외부와 연결된 공간, (2) 계단이나 엘리베이터와 같이 수직적인 연결이 있는 공간, 그리고 (3) 단절점(articulation point)들이 된다. 이 색칠된 노드들을 중심으로 색칠된 노드가 아닌 노드들을 병합한다. 그리고 색칠된 노드들 중에서 다른 색칠된 노드들과 최소 세 개의 에지로 연결된 노드를 결정점(decision point)으로 표시하고 이 결정점에 의해 닫히는 색칠된 노드들의 체인을 찾아서 색칠된 노드들을 하나의 노드로 병합한다. 결정점은 주로 복도나 계단과 같은 공간이 될 수 있다. 이렇게 해서 통합된 하나의 그래프는 다음 레벨에서의 기본 그래프가 되어 위의 과정을 반복한다.

이렇게 만들어진 계층 그래프는 주요 통로가 되는 공간을 중심으로 건물을 단순화시키기 때문에 내비게이션 같은 분야에서 효과적인 계층 구조이다 [6, 13].

2.3 연구 동기

앞서 설명한 공간 은폐를 위한 K-익명성을 활용하여 위치를 확장하는 기존의 방법들은 기본적으로 유클리드 공간을 전제로 한다. 사용자의 위치를 유클리디안 좌표를 기반으로 한 최소경계사각형으로

나타내어 확장시키거나 전체 공간을 일정한 크기의 격자로 나누어 위치를 은폐한다. 그러나 실외의 환경에서 제안된 기존의 공간 은폐 방법을 실내공간에 적용시키는 것은 적합하지 않다.

그림 1은 실내공간에서 기존의 방법대로 최소경계사각형을 이용할 경우 발생하는 문제의 예를 보여준다. 이동객체의 위치를 최소경계 사각형으로 나타내고, 익명화의 조건 K 가 3이라 가정한다. 실내공간에서 m_1 이 질의를 요청했을 때 K -익명성을 만족할 때까지 m_1 의 위치를 확장시킬 경우, 그림 1의 회색 영역과 같이 나타난다. 실제로 이동거리 상으로는 m_6 이 m_3 보다 가까운 위치에 있음에도 불구하고 유클리디안 거리로 선택되므로 인접성을 보장할 수 없다. 그리고 두 공간을 이동할 수 없는 벽을 포함하면서 이동객체의 움직임에 대한 연속성을 보장할 수 없는 위치의 형태가 된다. 이는 단순 사각형으로는 추상적인 위치로서의 표현이 부적합하다는 것을 의미한다. 또한 이 예제에서는 3명 중에 적어도 한사람은 Room102에 위치하고 있다는 추가적인 정보를 노출하게 된다. 그러므로 실내공간에 맞는 위치 표현 방법과 더불어 K -익명성을 만족하기 위한 위치 은폐 방법이 필요하다.

본 논문은 실내공간에서 K -익명화로 사용자의 위치를 은폐시키기 위해 계층 그래프를 이용하는 방법을 제안한다. 실내 공간의 계층 구조가 존재할 경우, 위치를 은폐할 때 사용자의 실제 위치를 나타내는 노드에서 실제 위치를 포함하고 있는 상위 노드를 선택함으로써 위치를 확장시킬 수 있다. 적어도 $K-1$ 명의 다른 사용자를 포함하도록 계층 구조에서 최소 레벨의 노드를 선택함으로써 위치를 은폐시키는 방법을 제안한다.

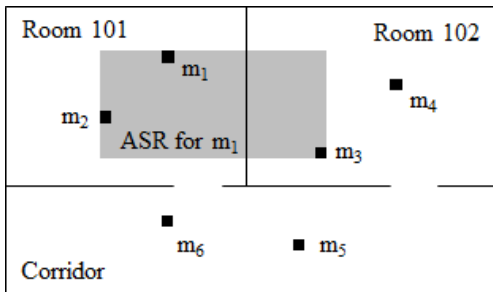


그림 1. 실내공간에서 기존 위치은폐 적용 예

3. 실내공간의 계층 구조와 위치

3.1 기호 공간

일반적으로 객체의 위치를 표현할 때 이동이 자유로운 공간에서는 좌표로 위치를 표현하는데 비해 이동의 제약이 있는 공간에서는 제약에 맞는 위치 체계를 사용한다. 실외공간은 유클리드 공간으로 표현되는데 반해 실내공간의 경우에는 벽, 천장, 바닥과 같은 구조물에 둘러싸인 제약공간이다. 실내에서 사람이 인지하는 방이나 복도, 계단과 같은 공간 단위로 위치를 표현하는 공간을 기호 공간이라 한다 [1]. 실내에서의 위치를 (2, 4, 3)의 3차원 좌표로 표현하지 않고 '1층 로비'라고 표현하는 것이 기호 공간의 한 예이다.

이러한 기호공간은 서로 겹치지 않는 공간들의 집합 N 과 그 공간들 간의 연결관계 집합 E 를 포함하는 그래프 $G=(N, E)$ 로 표현된다. 즉, 방과 복도와 같은 공간을 노드로 하고 방과 복도가 이동가능하다면 그 둘 사이의 연결 관계를 에지로 한다. 다른 층과 층을 연결하는 계단과 각 층의 복도와 연결관계도 에지로 표현한다. 다음은 기호 공간에서 위치를 표현하는 것을 정의한 것이다.

정의 1. 기호 공간에서의 위치

이동객체를 하나의 점으로 대표할 때, 이동객체 x 가 그래프 $G=(N, E)$, $n_i \in N$ 의 영역에 존재할 때 이동객체 x 의 위치는 다음의 식으로 정의한다.

$$Loc(G, x) = n_i$$

기호공간에서 단위 공간은 벽과 문들로 경계가 발생하는 방과 같은 공간뿐만 아니라 복도와 로비와 같이 경계가 구조물로 구분되지 않고 의미만으로 구분하여 분할할 수 있다. 한 층의 복도가 너무 클 경우, 동일한 의미를 가지는 복도로 할지라도 위치 정확도의 측면에서 부정확성을 줄이기 위해 공간을 일정 크기로 분할하여 단위 공간으로 표현 가능하다.

3.2 계층 구조와 위치

본 논문에서는 실내공간의 계층 구조를 계층 그래프를 통해 나타내고 있다. 실내 공간의 계층 그래프를 [13]에서 제시된 정의를 따라 다음과 같이 정의한다.

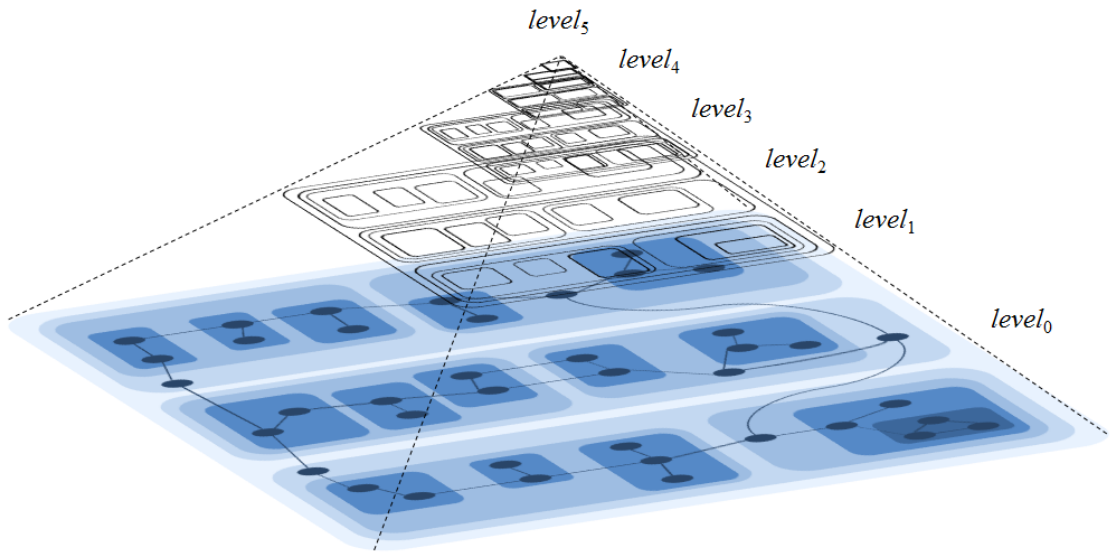


그림 2. 계층 그래프의 개념적 모델 예시

정의 2. 계층 그래프

$level_{l+1}$ 의 계층 그래프 H 는 $level_l$ 의 기본 그래프를 $k(> 1)$ 개의 연결된 노드들의 집합 $\{N_1, \dots, N_k\}$ 로 나눠 정의된다. 연결된 노드들의 집합인 각각의 $N_i \subseteq N$ 는 하위 그래프 $sub_i(G) = (N_i, E_i \subseteq E)$ 를 만든다. (이때, $E_i = \{(n_1, n_2) \in E | n_1, n_2 \in N_i\}$)

각각의 하위 그래프들은 계층 그래프 H 의 노드가 되며, G 에서 n_i, n_j 노드의 연결은 하위 그래프 $sub_i(G), sub_j(G)$ 의 연결이 되어 계층 그래프 H 의 에지가 된다.

그림 2는 3층 건물의 실내공간을 높이가 6인 계층 그래프의 예제이다. 그림의 가장 아래에 위치한 $level_0$ 그래프가 공간의 연결정보를 기반으로 한 기본 그래프이다. 검은 색의 타원이 계층 그래프의 단말 노드에 해당한다. 각 단말 노드들을 감싸고 있는 모서리가 둥근 사각형이 해당 노드를 포함하는 부모 노드가 된다. 이와 동일하게 사각형들을 감싸는 또 다른 사각형이 해당 사각형의 부모 노드이다.

실내공간의 노드로 위치를 표현하지만 노드들의 집합으로 위치를 표현할 수 있다. 예를 들어, 현재 나의 위치는 “422호”라고 표현할 수 있을 뿐 아니라 나의 위치는 “4층 A구역” 또는 “4층”이라고 좀더 추상적으로 위치를 표현 가능하다. 계층 그래프에서의 상위 레벨의 노드는 하위 레벨의 공간들의 집합으로 구성된다. “4층”은 “422호”를 포함하는 공

간으로서 계층 그래프의 관점에서 보았을 때 “422호”의 상위 노드가 된다. 다음은 계층 그래프를 통한 추상적인 위치 표현 방법을 정의한 것이다.

정의 3. 계층 그래프에서의 위치

이동객체 x 가 주어질 때, 기본 그래프 G_0 로 만들어진 계층 그래프 H 의 $l(0 \leq l \leq level)$ 레벨에서의 그래프를 $G_l = (N_l, E_l)$ 이라 할 때, l 레벨에서의 이동객체 x 의 위치는 다음과 같이 정의한다.

$$Loc(H, x, 0) = Loc(G_0, x)$$

$$Loc(H, x, l) = \{N_i | Loc(H, x, l-1) \in N_i, sub_i(G_l) = (N_i \subseteq N_l, E_i \subseteq E_l)\}$$

이동객체 x 의 위치는 l 레벨의 그래프에서 서로 겹치지 않는 노드들 중에 반드시 하나의 노드에만 공간적으로 속하게 되므로 위치는 해당 노드가 된다.

4. 계층 그래프를 이용한 K-익명화

4.1 K-익명화 서비스 구조 및 알고리즘

이 절에서는 K-익명화 서비스가 제공되는 흐름과 실내공간의 계층 그래프가 어떻게 K-익명성 은폐에 적용되는지 설명한다.

본 논문에서는 그림 3과 같이 클라이언트와 LBS 서버 사이에 위치 익명화를 수행하는 제 3의 신뢰

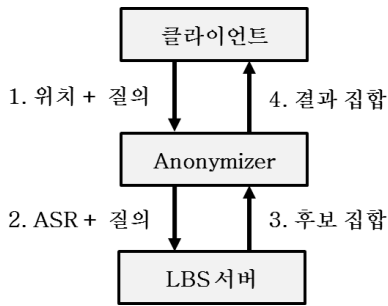


그림 3. 익명화 서비스 구조

할 만한 서버가 존재한다고 가정한다. 이러한 중앙 집중식의 구조는 이미 다수의 연구들에서 검증된 것이며, 이 구조에서는 각각의 LBS마다 자신의 익명화 수행 서버를 가지지 않아도 되며, 다양한 익명화 기법을 적용시킬 수 있다는 장점이 있다. Anonymizer를 포함한 전체 시스템의 개괄적인 K-익명화를 적용한 질의처리의 과정은 다음의 4단계로 진행된다.

1. 안전한 연결을 통해 클라이언트는 위치와 질의를 Anonymizer에게 전달한다.
2. Anonymizer는 클라이언트의 위치와 질의를 받아 K-익명화를 만족하는 ASR을 계산하여 LBS 서버에 ASR와 질의를 전달한다.
3. LBS 서버는 받은 ASR 기반 질의를 처리하여 후보 결과를 Anonymizer에게 전달한다.

Anonymizer는 사용자의 질의의 정확한 결과 집합을 찾아 클라이언트에게 반환한다.

Anonymizer에서 계층 그래프를 이용하여 K-익명화를 하는 구체적인 방법에 대해 위의 그림 4와 그림 5의 예를 통해 설명한다. 이해를 돕기 위해 사용한 층에 대한 평면도가 그림 4와 같이 실내공간 $S = \{R_1, R_2, R_3, R_4, R_5, C_1\}$ 이 구성된다고 가정한다. 그림 5의 왼쪽은 이를 계층 그래프로 만든 예이며, 그림 5의 오른쪽은 계층 그래프의 계층 구조를 쉽

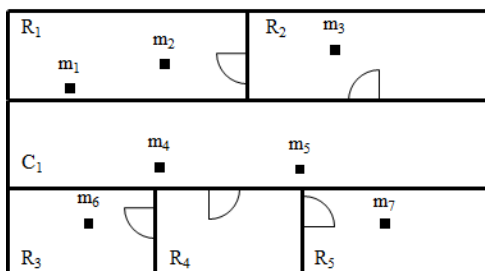


그림 4. 실내공간 평면도 및 이동객체 예제

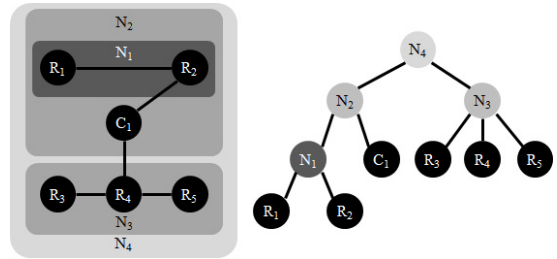


그림 5. 계층 그래프와 트리형태

게 이해하기 위해 노드들 계층 관계만을 트리로 표현한 것이다. 이 실내공간 S에 이동객체 $M = \{m_1, m_2, m_3, m_4, m_5, m_6, m_7\}$ 가 있으며 m_1 에서 $K=4$ 로 질의를 요청한다고 가정하자. m_1 는 R_1 에 위치하고 있으므로 R_1 노드에서 $K=4$ 명을 만족하는지 확인한다. R_1 에 존재하는 이동객체의 수는 $|\{m_1, m_2\}|=2$ 이므로 $K=4$ 조건을 만족하지 못한다. 계층 그래프에서 현재 노드의 부모 노드를 탐색한다. R_1 의 부모노드는 N_1 이고 N_1 노드는 $\{R_1, R_2\}$ 공간을 포함하는 노드이며, 이 공간에 있는 모든 이동객체의 수를 다시 K 값과 비교한다. N_1 노드에 있는 이동객체의 수는 $|\{m_1, m_2, m_3\}|=3$ 이므로 다시 N_1 의 부모 노드를 탐색한다. N_2 는 N_1 의 부모노드이고, $\{R_1, R_2, C_1\}$ 의 공간을 포함하는 노드이다. N_2 에 있는 이동객체의 수는 $|\{m_1, m_2, m_3, m_4, m_5\}|=5$ 이므로 $K=4$ 를 만족한다. 따라서 m_1 의 $K=4$ 를 만족하는 ASR는 N_2 가 된다.

다음의 그림 6은 앞서 설명한 방법대로 계층 그래프를 이용하여 K-익명성을 적용시키는 구체적인 과정을 알고리즘으로 보여준다. 익명화 조건 K와

Algorithm 1: K-anonymization

```

input: K // Condition of anonymization
input: Location // Location of requester
output: ASR // Anonymized Spatial Region
1: node ← FindNode(Location);
2: repeat
3:   k ← NumerOfMovingObject(node);
4:   if k ≥ K then
5:     break;
6:   end
7:   node ← node.parentNode;
8: until node = null;
9: return node;
    
```

그림 6. 계층 그래프의 K-익명화 알고리즘

질의 요청자의 위치를 입력받아 확장된 위치를 반환하는 알고리즘이다. 우선, 질의자의 위치를 나타내는 노드를 계층 그래프의 가장 낮은 레벨에서 찾는다. 찾은 노드로부터 자신의 상위 노드를 선택하면서 상위 노드에 존재하는 위치의 이동객체 수를 찾아 익명화 조건 K를 만족하면 해당 노드의 위치가 ASR이 된다.

4.2 K-익명화의 비용 분석

이 절에서는 앞서 살펴본 알고리즘을 적용하여 K-익명화를 수행하였을 때 예상되는 탐색 공간의 수와 수행 시간을 제시한다. 비용 모델을 간략화 하기 위해 몇 가지의 내용을 가정한다. 계층 그래프는 메인메모리 상에 로드되어 있다고 가정하고, 셀 공간인 단말노드에서 이동객체를 탐색하는 비용이 계층 그래프의 상위 노드들의 탐색 비용에 비해 아주 큰 것으로 간주한다. 계층 그래프 H는 노드 단말까지의 높이가 동일한 균형 트리이고, 평균 분기요소 bf, 레벨을 l이라 할 때, 레벨별 예상되는 탐색 단말의 노드의 수는

$$SearchLevelSpace(H,l) = bf^l$$

이다. 전체 실내공간의 수를 n이라 하고, 전체 이동객체의 수를 m이라 하고, 모든 공간에 균등하게 분포되었다고 가정할 때, K-익명성을 만족하기 위해

$$K \leq \frac{m \cdot bf^l}{n}$$

가 되어야 하므로 K에 따른 예상 탐색 레벨은 다음과 같다.

$$l \geq \left\lceil \log_{bf} \frac{m \cdot K}{n} \right\rceil$$

따라서 K에 따른 탐색 단말 노드의 수의 누적은

$$SearchAccumulativeSpace(H,K) = \sum_{l=0}^{\left\lceil \log_{bf} \frac{m \cdot K}{n} \right\rceil} bf^l$$

이다. 하지만 K-익명화 조건 탐색에서 각 레벨별로 이동객체 탐색 결과를 저장하고 상위 레벨로 올라가도 이전 탐색했던 공간을 다시 탐색하지 않으면 이를 누적할 필요가 없다. K-익명화의 탐색 단말 노드의 수는

$$SearchSpace(H,K) = bf^{\left\lceil \log_{bf} \frac{m \cdot K}{n} \right\rceil}$$

이다. 하나의 셀에 이동객체가 있는지를 탐색하는 평균 비용을 c이라 할 때, 계층 그래프 H의 K-익명화의 비용은 단말 노드의 수와 비례하고, 다음의 식으로 정리된다.

$$KAnonymityCost(H,K) = c \cdot bf^{\left\lceil \log_{bf} \frac{m \cdot K}{n} \right\rceil}$$

예를 들어, 단말 노드의 평균 탐색 비용 c=0.001 초 이고, 평균 분기요소가 bf=3, 이동객체가 m=1000개, 전체 공간의 수가 n=300개, K=25 이면 예상되는 탐색 비용은 다음과 같다.

$$KAnonymityCost(H_{bf=3},25) = 0.001 \cdot 3^{\left\lceil \log_3 \frac{300 \cdot 25}{1000} \right\rceil} = 0.009$$

이 때, 노드의 탐색 비용은 특정 범위에서 K의 증가에 따라 계단형으로 증가하게 된다. 즉, 앞의 예에서는 $10 \leq K \leq 30$ 범위에서 현재 노드에서 K를 만족하기 때문에 동일한 탐색 비용은 0.009초가 된다. 이러한 범위는 계층 그래프의 bf값에 의해 결정된다. $bf^i < n \cdot K/m \leq bf^{i+1}$ 인 범위에서 bf의 값이 노드의 수가 된다. 예를 들어 bf=2일 때, 구간 (1,2]구간 (2,4], 구간 (4,8]에서 각각 노드의 수가 2, 4, 8이 된다.

그림 7에서 bf 값이 2와 3인 경우일 때, x축을 $n \cdot K/m$ 로 하고 y축을 탐색 단말 노드의 수로 하

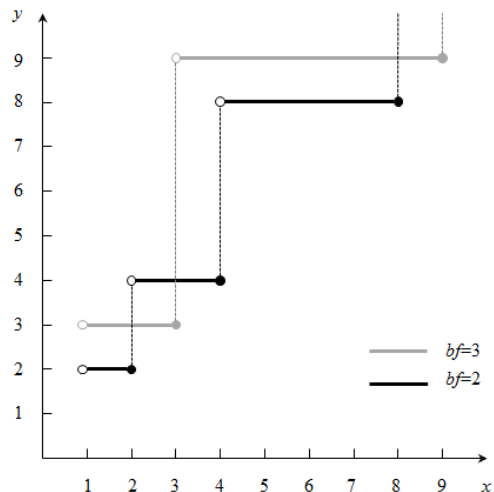


그림 7. bf의 2, 3의 탐색 노드 수 비교

여 비교한 그림이다. $bf=2$ 일 때는 검은색 그래프로 $bf=3$ 일 때는 회색 그래프로 나타난다. 예를 들어, $bf=2$ 이고 x 축 $n \cdot K/m$ 의 값이 (2,4] 구간 즉, 2초과 4이하이면 탐색 단말 노드의 수는 4가 된다. $bf=2,3$ 인 두 경우를 동시에 비교해 볼 경우, x 축 (3,8] 구간에서는 $bf=2$ 인 경우 탐색 단말 노드의 수가 8로 좋지만, (8,9] 구간에서는 $bf=3$ 인 경우가 노드 수 9로 더 적다. 따라서 모든 K 값에 대해 만족하는 bf 의 최적의 값은 존재하지 않는다. 다만 bf 값이 작을수록 구간의 범위가 작아지므로 상대적으로 K 의 범위를 만족할 확률이 높다. 예를 들어, x 축의 1에서 100까지 존재할 때, $bf=2$ 가 낮은 범위의 합은 69이고, $bf=3$ 가 낮은 범위의 합은 30이다. 계층 그래프를 생성할 때는 균형적이고 분기 요소의 수치가 낮아져야 K -익명화에 유리하다는 결과를 얻을 수 있다.

5. 결론

본 논문에서는 실내공간의 위치기반서비스에서 프라이버시 보호를 위해 계층 그래프를 이용한 K -익명성 위치 은폐 방법을 제안하였다. 기존의 위치 K -익명화 방법은 이동의 제약이 없는 유클리디안 거리를 기반으로 하였기 때문에 위치의 표현이 다른 실내공간에서 그대로 적용하기에 한계가 있었기 때문에 본 논문에서는 계층 그래프를 통해 이를 해결하였다.

위치 K -익명화에서 계층 그래프는 두 가지의 기능을 담당한다. 첫째로 계층 그래프는 실내공간에서 위치를 추상화하여 표현하기 위한 모델이다. 실외공간에서 위치를 추상화할 때 대표적으로 사용되는 것이 기하객체를 포함하는 경계사각형인데 이를 계층 그래프에서는 공간을 포함하는 상위 노드로 대신하여 표현한다. 둘째로 계층 그래프는 K -익명성을 만족하는 이동객체를 탐색하는 자료구조로 이용된다. 기존의 K -익명성 은폐 방법에서 쿼드-트리를 통해 이동객체를 탐색하는 것과 유사한 구조이다. 질의 요청자의 위치를 나타내는 노드에서부터 K -익명성을 만족하는 최소 레벨의 상위 노드를 ASR 로 선택함으로써 위치를 은폐시킨다.

계층 그래프의 형태에 따라 K -익명성 은폐의 성능이 달라지는데 익명화 조건 K 와 계층 그래프 속성에 따른 비용 모델을 제시하였다. 제안된 비용모

델에 따르면 균형적인 계층 그래프일 경우, 평균 분기요소가 작을수록 탐색할 노드의 수가 작아질 확률이 높아진다. 계층 그래프를 어떻게 생성하는가에 따라 위치 은폐의 기능이 달라질 수 있다. 따라서 위치 K -익명화를 위해 효율적인 계층 그래프를 만드는 것이 필요하다. 그리고 기존 K -익명화 연구들과 성능 및 보안에 대한 다양한 비교 실험을 통해 본 논문에서 제시한 계층 그래프를 이용한 K -익명화 방법이 실내공간에서 효율적이고 안정적인지를 검증하는 것이 향후 과제이다.

실내공간에서 특정 공간에 많은 인원이 분포하게 될 경우에 이 지역에 질의 요청이 들어오면 K -익명성을 만족하는 ASR 의 영역이 아주 좁아 질 수 있으며 실내공간의 특성이 분류될 수 있다. 따라서 [9, 15]에서 제시한 방법과 같이 실내공간에 대한 1-다양성을 반영하여 더 안정적인 K -익명화 방법에 대해 고려해 볼 수 있다.

참고 문헌

- [1] C. Becker and F. Durr, 2005, "On Location Models for Ubiquitous Computing," *Personal and Ubiquitous Computing*, vol. 9, no. 1, pp. 20-31.
- [2] B. Gedik, and L. Liu, 2005, "Location Privacy in Mobile Systems: A Personalized Anonymization Model," In *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems*, pp. 620-629.
- [3] A. Gkoulalas-Divanis, P. Kalnis and V.S. Verykios, 2010, "Providing K -anonymity in Location Based Services," *ACM SIGKDD Explorations*, vol. 12, no. 1, pp. 3-10.
- [4] M. Gruteser, and D. Grunwald, 2003, "Anonymous Usage of Location-Based Services through Spatial and Temporal Cloaking," In *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, pp. 31-42.
- [5] B. Hagedorn, M. Trapp, T. Glander, and J. Dollner, 2009, "Towards an Indoor Level-of-Detail Model for Route Visualization," In *Proceedings of the 10th International Conference*

- on Mobile Data Management: Systems, Services and Middleware, pp. 692-697.
- [6] H. Hu and D. Lee, 2004, "Semantic Location Modeling for Location Navigation in Mobile Environment," In Proceedings of the 5th IEEE International Conference on Mobile Data Management, pp. 52-61.
- [7] P. Kalnis, G. Ghinita, K. Mouratidis and D. Papadias, 2007, "Preventing Location-Based Identity Inference in Anonymous Spatial Queries," The IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 12, pp. 1719-1733.
- [8] K. Li, 2008, "Indoor Space: A New Notion of Space," In Proceedings of the 8th International Symposium on Web and Wireless Geographical Information Systems, pp. 1-3.
- [9] A. Machanavajjhala, D. Kifer, J. Gehrke and M. Venkitasubramaniam, 2007, "L-diversity: Privacy Beyond k-Anonymity," ACM Transactions on Knowledge Discovery from Data, vol. 1, no. 1, pp. 3:1-3:52.
- [10] M.F. Mokbel, C.Y. Chow, and W.G. Aref, 2006, "The New Casper: Query Processing for Location Services without Compromising Privacy," In Proceedings of the 32nd International Conference on Very Large Data Bases, pp. 763-774.
- [11] K. Richter, S. Winter and U. Ruetschi, 2009, "Constructing Hierarchical Representations of Indoor Spaces," In Proceedings of the 10th International Conference on Mobile Data Management: Systems, Services and Middleware, pp. 686-691.
- [12] H. Samet, 1990, The Design and Analysis of Spatial Data Structures, p. 510, Addison-Wesley Series in Computer Science.
- [13] E. Stoel, K. Schoder and H.J. Ohlbach, 2008, "Applying Hierarchical Graphs to Pedestrian Indoor Navigation," In Proceedings of the 16th ACM SIGSpatial International Conference on Advances in Geographic Information Systems, pp. 54.
- [14] 강혜영, 김준석, 황정래, 이기준, 2008, "셀룰러 공간에 존재하는 이동객체 궤적의 유사도 측정," 한국 GIS학회지, 제16권, 제3호, pp. 291-301.
- [15] 김지희, 이아름, 김용기, 엄정호, 장재우, 2008, "위치기반 서비스에서 개인 정보 보호를 위한 K-anonymity 및 L-diversity를 지원하는 Cloaking 기법," 한국공간정보시스템학회 논문지, 제10권, 제4호, pp. 1-10.
- [16] 엄정호, 김지희, 장재우, 2009, "위치기반 서비스에서 개인 정보 보호를 위한 그리드를 이용한 Cloaking 영역 생성 알고리즘," 한국공간정보시스템학회 논문지, 제11권, 제2호, pp. 151-161.

논문접수 : 2012.05.11

수정일 : 2012.08.07

심사완료 : 2012.08.14



김 준 석

2005년 부산대학교 정보컴퓨터공학부
학사

2007년 부산대학교 컴퓨터공학과 석
사

2007년~현재 부산대학교 컴퓨터공학
과 박사과정

관심분야는 Indoor GIS, 시공간데이터베이스, 3차원 공
간데이터 모델링, 텔레메틱스



이 기 준

1984년 서울대학교 전자계산학과 학
사

1986년 서울대학교 전자계산학과 석
사

1992년 프랑스로립응용과학원 전자계
산학과 전산학 박사

1992년~현재 부산대학교 정보컴퓨터공학부 교수
관심분야는 공간 및 시공간데이터베이스, GIS, 유비쿼
터스 컴퓨팅