위치기반서비스의 개인정보보호를 위해 Dummy를 이용한 Cloaking 영역 설계

정회원 김 주 용*, 정 은 희**°, 이 병 관***

A Design of Cloaking Region using Dummy for Privacy Information Protection on Location-Based Services

Ju-Yung Kim*, Eun-Hee Jeong**°, Byung-Kwan Lee*** *Regular Members*

요 약

본 논문에서 제안한 Dummy를 이용한 Cloaking 영역 설정 알고리즘은 기존의 Privacy Grid의 색인 구조에 건물 그룹화 항목을 추가시켜 Privacy Grid의 여러 셀에 걸쳐져 있는 건물을 중복으로 카운트하던 문제를 해결하였고, 각 건물 모서리에 인접한 셀을 검색한 후 K값을 증가시킴으로써 기존의 GBC(Grid-Based Cloaking) 기법의 최소 Cloaking 영역 설정으로 인한 개인 위치 정보 보호 유출 문제점을 해결 하였다. 또한, 본 논문에서는 Cloaking 영역을 확장하기 전에 Dummy K 값으로 Privacy Grid와 GBC 보다 작은 Cloaking 영역을 설정할 수 있어 K값 검색시간을 단축시키고, Dummy K를 이용하여 사용자의 위치정보보호를 더 강화시킬 수 있다

Key Words : Cloaking, Dummy, LBS, K-Anonymity, L-Diversity

ABSTRACT

The setting algorithm of cloaking region using dummy which is proposed in this paper solves the problem which counts the building with duplication that exists in several cells of Privacy Grid by adding the building grouping item to the index structure of the existing Privacy Grid, and by increasing K value after searching the contiguous cells in the corner of each building, the exposure problem of private location information due to the minimum cloaking region setting of the existing GBC is solved. In addition, this paper reduces the searching time of K value by setting smaller cloaking region than Privacy Grid and GBC with dummy K before expanding cloaking region and strengthens the location information protection of users using dummy K.

Ⅰ. 서 론

최근 이동통신 기술의 발달과 함께 이동단말의 위치를 파악하고, 이를 기반으로 다양한 서비스를 제공하고자하는 위치기반서비스(이하 LBS : Location based Service)의 수요가 증가함에 따라 많은 국내외 민간기업, 공공기관, 그리고 지자체 등에서는 이 위치정보를 가공하여 민간과 공공부문에 서비스할 다양한 콘텐츠를 개발하고 있다. 특히, 스마트폰의 보급이 급격히 확산되고 구글맵 등 주요 서비스가 광고와 연계해 무료로 제공됨에 따라 2013년경 스마트폰 이용자의 80%가 LBS를 이용할 것으로 예상된다. 또한 전 세계 모바일 LBS 시장은 2009년 21억 달러에서 2015년 183억 달러로 연평균 43% 씩 고성장할 전망이다[1].

그러나 LBS는 서비스를 요청하는 사용자의 정확한 위치 정보를 데이터베이스 서버에 보내기 때문에 서

* 관동대학교 전자계산공학과 (jjeleun@nate.com), ** 강원대학교 지역경제학과 (jeongeh@kangwon.ac.kr), (° : 교신저자)
*** 관동대학교 컴퓨터학과 (bklee@kd.ac.kr)
논문번호 : KICS2011-05-214, 접수일자 : 2011년 5월 12일, 최종논문접수일자 : 2011년 6월 10일

비스 이용자들이 어떤 장소에 자주 방문하는지, 또한 이러한 방문이 어떤 시간대에 주로 이루어지는지를 파악하여, 개인 정보를 획득할 수 있으므로, 사용자의 개인 정보가 노출될 수 있는 취약성이 높다고 할 수 있다.

실제로, 위치 기반 서비스를 이용한 스토킹 피해 사례가 빈번히 발생하고 있으며, 이러한 위치정보서비스의 이용 확대와 더불어 개인의 위치정보 프라이버시의 침해에 대한 위험성 또한 증가하고 있다. 따라서 모바일 사용자의 안전한 위치기반 서비스 사용을 위해서는 개인 정보 보호 방법이 요구된다.

본 논문에서는 이러한 개인정보보호를 위해 K-anonymity와 L-diversity를 이용한 Privacy Grid Cloaking 기법과 Grid Based Cloaking 기법을 살펴보고, 각 Cloaking 기법의 문제점을 분석하여, 그 문제점에 대한 해결방안으로 Dummy를 이용한 Cloaking기법을 설계해 위치기반서비스에서 개인정보를 보호하고자 한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 개인정보 보호 방법 연구 중의 Cloaking 기법을 살펴보고, 기존의 Cloaking 기법인 Privacy Grid Cloaking 기법과 Grid Based Cloaking 기법의 문제점들을 분석해본다. 그리고 3장에서는 본 논문에서 제안하는 Dummy를 이용한 Cloaking 영역 설계 기법을 기술한다. 4장에서는 3장에서 제안한 Dummy를 이용한 Cloaking 기법을 실험 및 평가하고, 마지막으로 5장에서 결론을 맺는다.

Ⅱ. 관련 연구

위치기반서비스에서 개인 정보유출을 방지하기 위해 송수신되는 메시지 암호화, 사용자 개인 정보 제공에 대한 정책, 사용자의 익명성을 보장하는 방법이 연구되고 있으며[2,3,4], 사용자의 정보 노출을 방지하기 위해 K-anonymity나 L-diversity와 같은 익명화 모델이 제안되었다[5,6]. 본 절에서는 K-anonymity와 L-diversity를 이용한 사용자의 익명성을 보장하는 Cloaking 기법에 대해 살펴본다.

2.1 중앙집중방식의 Cloaking 기법

Cloaking기법은 개인정보를 보호하는 방법 중의 하나로, 사용자가 위치정보에 대한 서비스를 요청하면, Anonymizer 서버는 사용자의 좌표 정보가 아닌 K-anonymity를 만족하면서 최소한의 넓이를 가지는 질의 영역인 Cloaking을 생성하여 LBS 서버로 전송한다. 이때, LBS 서버는 Cloaking 영역을 바탕으로 요청된 질의를 처리하여 질의수행결과를 Anonymizer 서버에 전송하면, Anonymizer 서버는 질의 처리된 결과를 저장된 사용자의 좌표정보를 바탕으로 필터링하여 정확한 정보는 사용자에게 전송하지만, LBS 서버에는 사용자의 위치를 숨김으로써 사용자의 위치정보를 보호하게 된다[1]. 그런데 K-anonymity만을 만족하는 Cloaking 기법은 Cloaking 영역이 동일한 건물과 같은 한 장소 내에 생성될 경우, 사용자의 위치가 쉽게 추정될 수 있는 문제점을 가지고 있다. 이와 같은 문제점을 해결하기 위하여 Cloaking 영역 내에 L개의 다른 장소를 포함시키는 L-diversity가 고려되었고, 이는 K-anonymity와는 보완적으로 사용된다.

그림 1은 중앙집중방식 Cloaking기법의 Anonymizer 서버가 Cloaking 영역을 설정해 LBS 서버에 전송하는 처리절차를 간단히 설명한 것이다[7]. 중앙집중방식 Cloaking 기법은 영역 설정시간은 빠르지만, Anonymizer 서버의 병목현상으로 인한 성능저하 문제 및 사용자의 위치 정보를 저장하고 있는 Anonymizer 서버의 보안 위협 문제가 존재한다.

이러한 중앙 집중 방식의 문제점을 해결하기 위해 분산 방식 Cloaking 기법이 있는데, 이 기법은 Cloaking 영역 설정시 Cloaking 영역을 요구하는 사용자가 다른 사용자들과 통신하여 직접 Cloaking 영역을 설정한다. 하지만, 분산 방식 Cloaking 기법은 분산 구조로 높은 신뢰성을 제공하지만, 사용자들 간의 통신비용이 들고, Cloaking 영역 설정 속도가 중앙집중 방식보다 느리다는 단점이 있다.

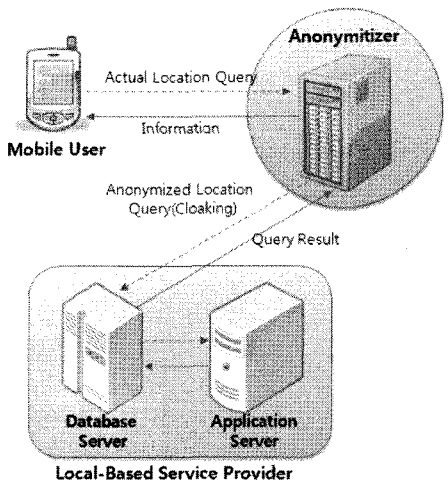


그림 1. 중앙 집중 방식의 Cloaking 기법

2.2 K-anonymity와 L-diversity의 Cloaking 기법

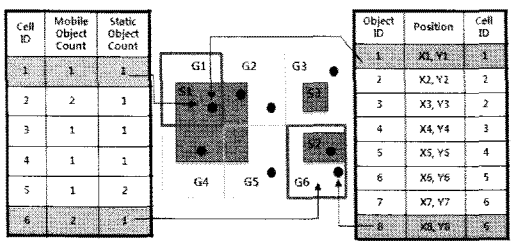
2.2.1 Privacy Grid Cloaking 기법

K-anonymity와 L-diversity를 동시에 고려한 Cloaking 기법으로는 B. Bamba와 L. Liu가 제안한 Privacy Grid Cloaking 기법이 있다[8,9]. Privacy Grid는 다단계 그리드를 사용하고, 그리드 셀을 이용해 Cloaking 영역을 설정하므로, 전체 영역을 같은 크기의 그리드 셀로 나누고, 각 그리드 셀은 셀 내에 위치한 사용자의 수, 위치 그리고 건물 개수를 저장한 색인 구조인 Cell Object Count Map을 사용한다. Privacy Grid는 셀의 확장 방법에 따라 표 1과 같이 분류되는데[8,9], 그리드 셀을 이용하여 Cloaking 영역을 생성하기 때문에 Cloaking 시간이 매우 빠른 장점을 가진다.

그림 2는 Privacy Grid의 Cell Object Count Map, Privacy Grid Index, 그리고 Cloaking 영역 내에 있는 건물(Static Object)의 개수, 사용자(Mobile Object)의 수를 설명한 것으로[2,8], 그림 2-(a)는 그리드의 셀에 포함되어 있는 모바일 개수와 건물의 개수를 저장하는 Cell Object Count Map의 구조이고, 그림 2-(b)는 각각의 모바일의 위치와 모바일이 속해 있는 그리드의 셀 번지를 저장하는 Privacy Grid Index의 구조이다.

표 1. Privacy Grid Cloaking 영역 설정 방법

셀 확장 방법	설 명
Quad-Grid	전체 영역을 4등분하여 각각을 그리드 셀로 Cloaking 영역으로 설정
Bottom-Up	질의를 요청한 사용자가 위치한 셀을 기준으로 셀을 확장하면서 Cloaking 영역을 설정
Top-Down	사용자가 정의한 최대 Cloaking 영역 크기에서 셀을 감소시키면서 Cloaking 영역을 설정
Hybrid	사용자가 요구한 K-anonymity와 최대 Cloaking 영역의 크기에 따라 Bottom- Up과 Top-Down 방식 중에서 하나를 선택하여 Cloaking 영역 설정



(a)Cell Object Count Map구조 (b)Privacy Grid Index구조

그림 2. Privacy Grid의 데이터 구조

Privacy Grid Cloaking기법의 첫 번째 문제점은 여러 그리드 셀에 걸쳐져 있는 하나의 건물을 여러 그리드 셀의 Static Object에 중복해서 카운트하는 문제점을 가지고 있다.

예를 들어, 그림 2-(a)에서 건물 S1이 그리드 셀 G1, G2, G4, G5에 걸쳐져서 위치할 경우, 그리드 셀 G1, G2, G4, G5의 Static object count가 모두 하나씩 증가된다. 이것은 건물 4개가 포함하는 Cloaking 영역을 요구 했을 때, Cell Object Count Map을 참조하여 Cloaking 영역이 G1, G2, G4, G5의 셀이 선택되었다면, 건물 4개를 포함한 것처럼 보이지만, 실제로는 하나의 건물만이 포함된 Cloaking 영역이 설정되므로 건물 4개를 포함하는 Cloaking 영역 설정 요구에 위배된다. 이러한 문제점을 해결하기 위해, 여러 셀에 걸쳐져 있는 Static Object가 중복되어 카운트 되는 것을 해결해야 한다. 본 논문에서는 여러 셀에 걸쳐져 있는 동일 건물을 그룹화 하여 이 그룹 정보를 데이터 구조에 추가하여 Static Object의 중복 카운트 문제를 해결하고자 한다.

두 번째 문제점으로는 L-diversity와 K-anonymity를 만족하는 Cloaking 영역을 설정하였지만, Cloaking 영역 내에 단순히 건물만 포함시키고 건물 내에 위치한 타사용자의 수를 고려하지 않아 질의를 요청한 사용자의 위치가 노출될 확률이 높을 수 있다. 그림 3의 점선으로 표시된 Cloaking 영역은 사용자의 요구를 통해 하나의 건물과 3명의 사용자가 위치한 모습을 두 개의 셀로 나타낸 것이다. 사용자의 요구에는 충족하였지만 실제 건물 S2에는 타사용자가 위치하지 않기 때문에 서비스 사용자가 도로상에 위치했다는 것을 쉽게 추측할 수 있을 것이다. 이는 L-diversity의 특성을 충분히 고려하지 못한 것으로 본 논문에서는 Cloaking 영역에 포함된 건물 내에 사용자의 분포 상황에 대한 전제조건을 제시함으로써 한명 이상의 사용자가 위치한 건물을 Cloaking영역을 포함시키도록 하여 이 문제를 해결하고자 한다.

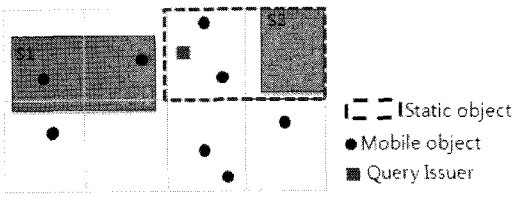


그림 3. Cloaking 영역 설정

2.2.2 Grid-based Cloaking 기법

Privacy Grid Cloaking 기법의 두 가지 문제점을 해결하기위해 제안된 Cloaking 기법으로 Grid-based Cloaking(GBC) 기법을[2]에서 제안하였다. 하지만, GBC 기법은 L-diversity를 만족하는 최소경계사각형을 임시 Cloaking 영역으로 설정하고, 이 임시 Cloaking 영역이 K-anonymity를 만족한다면 사용자의 위치 K-a보호하는 Cloaking 영역으로 설정된다. 그런데, 그림 4처럼 4개의 건물과 8명의 사용자를 포함하는 Cloaking 영역이 설정되었지만, 각 건물의 모서리로 영역이 설정된다면 해당 사용자가 건물외부 도로에 있다는 것을 쉽게 유추할 수 있기 때문에 L-diversity 특성을 고려하지 못한 것으로 건물의 모서리에 치우치지 않는 Cloaking 영역 설정방법이 필요하다.

두 번째 문제점으로는 임시 Cloaking 영역이 건물의 위치에 따라 Cloaking 영역 자체가 불필요하게 커질 수 있는 문제점을 들 수 있다. GBC 기법은 Cloaking 영역을 한 셀씩 확대하여 L값이 3이상에서 최소 경계 Cloaking 영역을 반환하지만 실제로 그림 5와 같은 건물구조일 경우에는 L값이 4인 Cloaking 영역을 임시 Cloaking 영역으로 반환하는 것을 확인

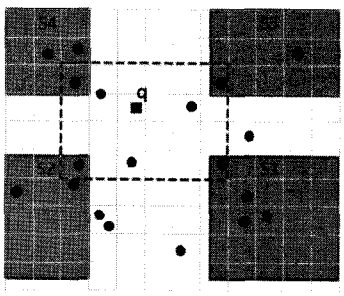


그림 4. 최소 Cloaking 영역 설정

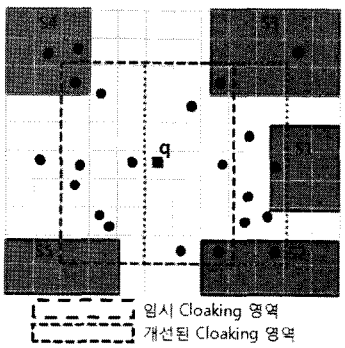


그림 5. 건물위치에 따른 잘못된 Cloaking 영역 설계

할 수 있다. 그러나 S1, S2, S3을 포함하는 Cloaking 영역이 좀 더 최소화 되는 것을 확인 할 수 있다. 본 논문에서는 L-diversity와 K-anonymity를 만족하는 임시 Cloaking 영역 설정 후, 각 건물간의 거리를 측정하여 사용자의 위치와 가장 근접하면서 각각의 건물간의 거리도 근접한 건물만이 포함되도록 Cloaking 영역을 보정하여 Cloaking 영역이 불필요하게 커지는 것을 문제점을 해결하고자 한다.

세 번째 문제점은 L-diversity는 만족하는 임시 Cloaking 영역을 지정한 후 임시 Cloaking 영역이 K-anonymity를 만족하지 않는 경우, K-anonymity를 만족할 때 까지 확장 Cloaking 알고리즘을 통해 영역을 확장하는데, 이로 인해 Cloaking 영역 설정 속도가 느리다는 단점을 가지고 있다. 본 논문에서는 K-anonymity를 만족하기 위해 영역을 확장하기 전에 Dummy K를 생성하여 K-anonymity를 만족시킴으로써 Cloaking 영역 설정 속도를 향상시키고, 사용자의 위치정보보호를 향상시키고자 한다.

Ⅲ. Dummy Cloaking 영역 생성 알고리즘 설계

본 논문에서 제안하는 Dummy Cloaking 영역 생성 알고리즘은 Cloaking 영역설정을 하기 위하여 기존 연구방법인 Grid를 이용하였으며 먼저 L-diversity 값을 충족시킨 후 K-anonymity를 고려하는 방법을 선택하였다. 그리고 각 그리드 셀을 Privacy Grid의 Cell Object Count Map, Privacy Grid Index와 같은 데이터 구조를 이용하여 각 셀에 포함된 사용자와 건물의 위치를 저장시키는 방법을 사용하였다.

또한, L-diversity를 만족하는 임시 Cloaking 영역이 K값을 만족하지 못하면 Cloaking 영역 확장 알고리즘을 실행하는데, 이때 Dummy K를 생성할 조건이 맞는지 검사한 후, 생성 조건에 만족하는 상황이면 Dummy K를 먼저 생성하도록 하여 확장 알고리즘을 좀 더 빠르게 종료하도록 설계하여 기존의 Privacy Grid Cloaking 기법과 GBC 기법의 문제점들을 해결하고자 한다.

그림 6은 Dummy Cloaking 기법의 전체적인 흐름을 설명한 것이다.

3.1 색인구조 설계

Privacy Grid는 여러 그리드 셀에 걸쳐져 있는 하나의 건물을 각 셀의 Static Object에 각각 카운트하여 중복적으로 포함시키는 문제점이 있었다. 이와 같은 문제점을 해결하기 위해 본 논문에서는 여러 셀에 걸

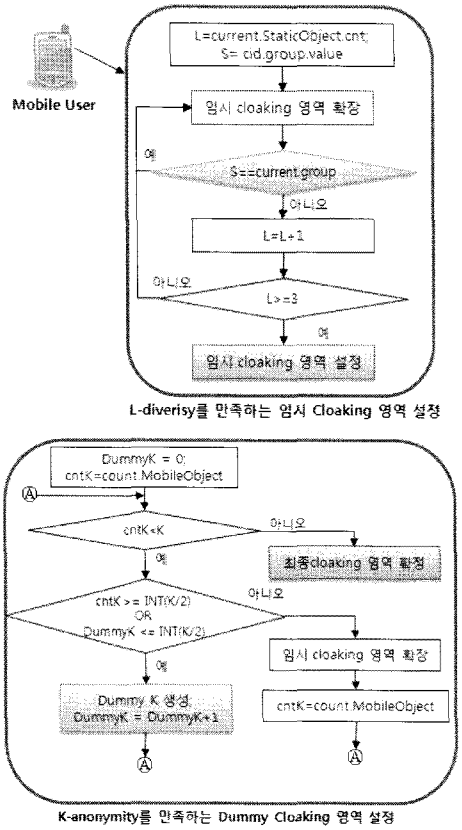
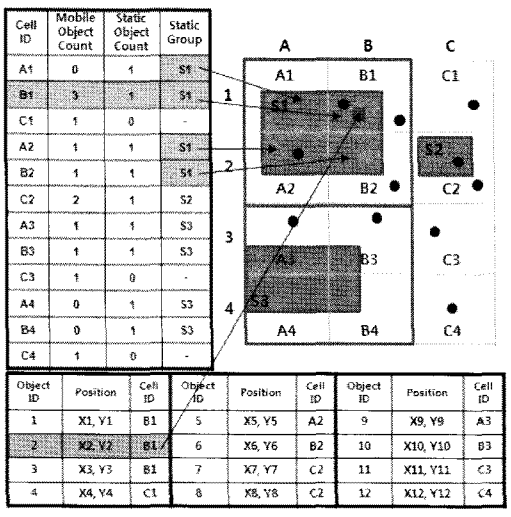


그림 6. Dummy를 이용한 Cloaking 영역 설계 흐름도

쳐져 있던 하나의 건물을 그룹화 시킨 데이터 값을 색인구조에 포함시켜 설계하였다. 그림 7은 그룹 정보가 포함된 색인구조를 설명한 것이다.



Cell ID	Mobile Object Count	Static Object Count	Static Group
A1	0	1	S1
B1	3	1	S1
C1	1	0	-
A2	1	1	S1
B2	1	1	S4
C2	2	1	S2
A3	1	1	S3
B3	1	1	S3
C3	1	0	-
A4	0	1	S3
B4	0	1	S3
C4	1	0	-

Object ID	Position	Cell ID	Object ID	Position	Cell ID	Object ID	Position	Cell ID
1	X1, Y1	B1	5	X5, Y5	A2	9	X9, Y9	A3
2	X2, Y2	B1	6	X6, Y6	B2	10	X10, Y10	B3
3	X3, Y3	B1	7	X7, Y7	C2	11	X11, Y11	C3
4	X4, Y4	C1	8	X8, Y8	C2	12	X12, Y12	C4

그림 7. 그리드 셀의 데이터 정보에 관한 색인구조

예를 들어 그림 7에서처럼 A1, B1, A2, B2 셀에 하나의 건물 S1이 걸쳐져 있는 경우, S1이라는 그룹으로 묶은 후, 그 정보를 색인구조에 저장한다.

S1={A1,B1,A2,B2}, S2={C2},
S3={A3,B3,A4,B4}

이와 같은 건물그룹 데이터를 통해 임시 Cloaking 영역 설정시 같은 그룹을 가진 건물은 하나의 건물로 인식하여 임시 Cloaking영역을 지정한 L값을 충족할 수 있게끔 확대 시킨다.

3.2 Dummy Cloaking 영역 생성 알고리즘 설계

Dummy Cloaking 영역 생성 알고리즘은 L-diversity 설정 후 K-anonymity를 설정하는 것으로 설계하였으며, 5개의 수행단계로 구성된다.

3.2.1 수행단계 1. 사용자가 요구한 L개의 건물 선택($L \geq 3$)

사용자가 지정한 값이 L=3일 경우 그림 8과 같이 사용자를 기준으로 한 셀씩 Cloaking영역을 확장 시키면 1단계 임시 Cloaking영역 내에는 3개의 건물을 포함하는 Cloaking 영역으로 설정되는 것처럼 보일 수 있다. 하지만 D3과 E3은 색인구조의 static group 값이 S4라는 동일한 값을 가지므로 동일 건물로 간주하고 Cloaking영역을 한 셀씩 확대하여 2단계 임시 Cloaking 영역으로 확장한다. 이때 L값이 3이상이면 Cloaking영역 확장을 종료한다.

본 논문에서는 여러 그리드 셀에 한 건물이 걸쳐져 있는 경우, 건물이 걸쳐져 있는 셀을 그룹화 시켜 그 그룹 값을 저장하도록 색인구조를 설계하여, Cloaking 영역에서 건물의 수를 카운트할 때 여러 그리드 셀에 걸쳐서 있는 건물을 중복적으로 카운트하는 Privacy Grid의 문제점을 해결하였다.

Cell ID	Mobile Object Count	Static Object Count	Static Group
:	:	:	:
B2	1	1	S4
C2	1	1	S4
:	:	:	:
C3	1	1	S4
D3	1	1	S4
E3	0	1	S4
:	:	:	:
H3	1	1	S3
:	:	:	:

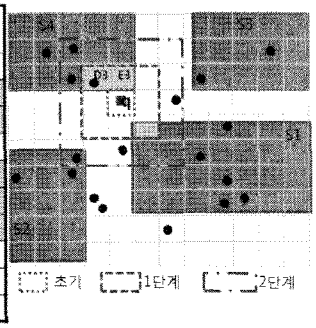


그림 8. Cloaking 영역 확장

그림 9는 사용자가 요구한 L개의 건물을 포함한 Cloaking 영역을 설정하는 알고리즘을 설명한 것이다.

```
Algorithm 1. L_Cloaking Algorithm
1: Function L_Cloaking(Q, L){
2:    cid = getCell(Q);
3:    s = cid.group.value;
4:    tempRegion = cid.MobileObject;
5:    cntL = current.StaticObject.cnt;
6:    while(cntL<3){
7:       tempRegion
           = ExpansionOfTempCloakingRegion(cid);
8:       add(glist, tempRegion.group.value);
9:       cntL = compare(glist);
10:   }
11:   return tempRegion;
12: }
```

그림 9. 사용자가 요구한 L개 건물을 포함한 Cloaking 영역 생성 알고리즘

3.2.2 수행단계 2. 임시 Cloaking 영역 설정

수행단계1에서 L값을 만족하는 임시 Cloaking 영역 설계를 종료했을 때, 그림 10-(a)과 같이 점선의 Cloaking 영역으로 설정된다면 Cloaking 영역의 크기가 불필요하게 크게 설정된 것이므로 임시 Cloaking 영역을 축소시키기 위한 Cloaking 영역설정 방법이 필요로 하다.

본 논문에서는 수행단계1을 통하여 L값이 만족하는 임시 Cloaking 영역이 설정되면, L값을 가지는 셀 정보를 통하여 사용자 Q와 셀 정보를 통하여 인접한 L값을 가지는 셀을 검색한다. 그리고 난 후, 사용자 Q를 통한 각 건물정보를 가지는 임의의 셀 값을 통해서 건물간의 거리를 측정을 한 후 건물간의 거리가 가장 작은 셀 값을 포함하는 최소 경계 임시 Cloaking 영역을 설정 한 후 수행단계3을 진행한다.

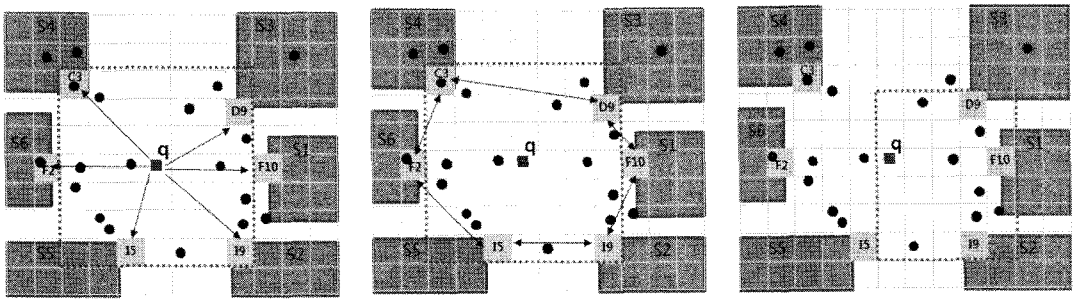
복잡한 도시에서는 수행단계1의 Cloaking 영역 설정 방법 보다는 수행단계2를 통해서 임시 Cloaking 영역 설계를 하는 것이 더 바람직하다. 그러나 건물의 수가 적은 소규모 도시와 같은 경우는 오히려 수행단계1을 수행함으로써 수행시간을 단축시키는 것이 더욱 효과적이라 본다.

본 논문에서는 이와 같이 분류 해 놓은 이유는 수행단계1을 실행함으로써 일단 L개 이상의 건물을 탐색한 후 수행단계2에서는 L개 건물에 인접해 있는 또 다른 건물을 통해 Cloaking 영역의 크기가 축소 할 수 있는 상황인지를 확인함으로서 최소 경계 임시 Cloaking 영역을 설정하기 위해서이다.

3.2.3 수행단계 3. 사용자가 요구한 K값이 임시 Cloaking 영역에 만족할 경우

수행단계3에서는 L-diversity를 만족한 임시 Cloaking 영역이 K값을 만족한다면, 해당 임시 Cloaking 영역을 반환하여 알고리즘을 종료하는데, GBC 기법에서처럼 각 건물의 모서리에 Cloaking 영역이 설정되는 최소 Cloaking 영역 설정 문제점이 발생할 수 있다.

예를 들어, 그림 11의 C3, H3, C6, H6의 건물 모서리를 포함하는 최소 Cloaking 영역은 사용자가 요구한 값이 L=3이고, K=12일 경우 요구하는 조건에 만족하므로 최소 Cloaking 영역이 최종 Cloaking 영역으로 확정된다면 L-diversity의 특성을 고려하지 않은 Cloaking 영역이 설정되는 것이며, GBC 기법의 최소 Cloaking 영역 설정 문제점이 발생한다. 하지만, 본 논문에서는 그림 11의 C3, H3, C6, H6의 건물 모서리만을 포함하는 최소 Cloaking 영역이 임시 Cloaking 영역으로 설정된다면, 건물에 포함되어 있는 K값을 증가 시키는 방법을 통해 모서리에 걸쳐져



(a) Q와 인접한 셀 검색 (b) 셀과 셀의 거리 측정 (c) 거리측정을 통한 임시 Cloaking 영역

그림 10. 임시 Cloaking 영역 설정

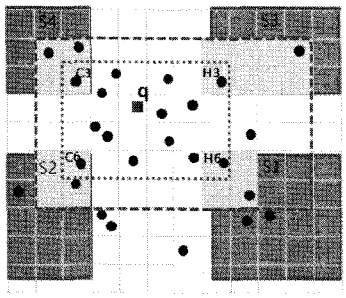


그림 11. L-diversity와 K-anonymity를 만족하는 최소 Cloaking 영역 확대

서 임시 Cloaking영역이 설정되는 최악의 경우를 회피한다. 즉, 각 모서리에 해당 사용자가 없을 경우 K ≥1값이 충족되도록 그림 11과 같이 임시 Cloaking 영역을 확대하여 최소 Cloaking 영역이 최종 Cloaking 영역으로 확정되는 것을 방지함으로써 GBC 기법의 최소 Cloaking 영역 설정문제를 해결하고, 개인 정보 보호를 더욱 더 강화시킨다.

그림 12는 L-diversity와 K-anonymity를 만족하는 Cloaking 영역을 설정하는 알고리즘을 설명한 것으로, LK_Cloaking 알고리즘은 건물의 모서리에 포함된 K값이 1이하이면 Cloaking 영역을 확장하도록 설계한 알고리즘이다.

```
Algorithm 2. LK_Cloaking Algorithm
1: Function LK_Cloaking(Q, K, L, tempRegion){
2:    cntK = CountK(tempRegion)
3:    while(cntK < K){
4:       tempRegion =
           ExpansionOfTemp
           _CloakingRegion(tempRegion);
5:       cntK = CountK(tempRegion);
6:       glist = FindGroup(tempRegion);
7:    }
8:    while(CountK(glist.group) <= 1){
9:       tempRegion =
           ExpansionOfTemp
           _CloakingRegion(tempRegion);
10:      glist = FindGroup(tempRegion)
11:   }
12:   return tempRegion;
13: }
```

그림 12. L값과 K값을 만족하는 Cloaking 영역 생성 알고리즘

3.2.4 수행단계 4. 사용자가 요구한 K값 선택($K' \geq K$)

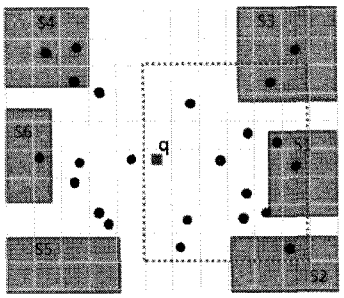
수행단계4에서는 사용자가 요구한 K값을 만족 못한다면 확장 알고리즘을 통하여 K값을 증가 시키는 단계로써 먼저 임시 Cloaking 영역 내에 K값을 확인한 후 사용자가 요구한 K값의 50%이상이면 Dummy를 생성하여 Cloaking 영역 내에 랜덤한 위치에 생성한다. Dummy를 생산하여도 악의적인 공격자가 모니터링 한다고 하더라도 실제로 존재하는 K값이 50%이상이면 공격자가 해당 사용자를 찾기는 어려울 것이다.

그림 13에서는 사용자가 요구한 K값이 17이면 그림 13-(a)은 K값이 13이다. 여기서 K값을 15이상 맞추기 위해서는 임시 Cloaking 영역을 확장알고리즘을 통하여 확장시켜 나가야 하지만 GBC 알고리즘과 같이 속도가 느리다는 단점을 가지고 있다. 그렇기 때문에 Dummy를 활용해 랜덤적으로 가상의 사용자를 Cloaking 영역 내에 요구한 K값에 맞춰 생성한다.

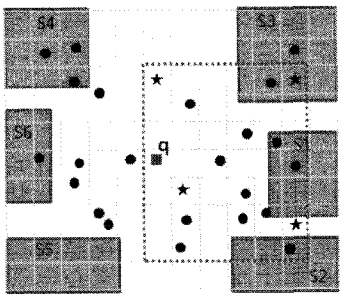
그림 13-(b)은 Dummy를 생성한 예를 나타낸 것이다.

Dummy를 사용하여 사용자가 요구한 K값을 만족하면 알고리즘을 종료 시키며 사용자가 요구한 K값이 임시 Cloaking 영역 내에 있는 K값이 50%를 넘지 않는 경우 수행단계5로 넘어간다.

그림 14는 Dummy를 생성하는 알고리즘을 설명한 것이다.



(a) Dummy 생성 전



(b) Dummy 생성 후

■User ●Another User ★Dummy Cloaking Zone

그림 13. Dummy를 활용한 Cloaking 영역

```
Algorithm 3. Dummy_Cloaking Algorithm
1: Function Dummy_Cloaking(Q, K, L, tempRegion){
2:    cntK = CountK(tempRegion)
3:    if (cntK >= Int(K/2)){
4:       while(cntK < K || DummyK <= Int(K/2) ){
5:          Dummy = CreateDummy(tempRegion);
6:          AddObject(CellObjectMap, Dummy);
7:          DummyK = DummyK + 1;
8:          if (DummyCntK >= Int(K/2)) break;
9:          cntK = CountK(tempRegion) ;
10       }
11:   }
12: return tempRegion;
13: }
```

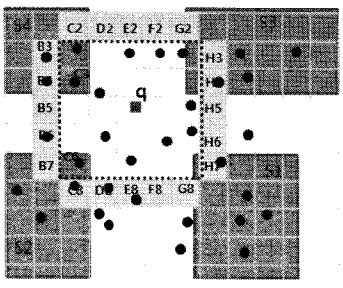
그림 14. Dummy Cloaking 영역 생성 알고리즘

3.2.5 수행단계 5. 확장 Cloaking영역 설정

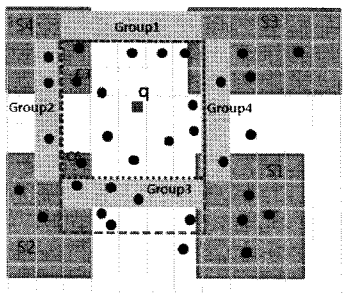
사용자가 요구한 K의 값이 K=17일 경우 임시 Cloaking 영역에서 가로와 세로를 한 셀 증가한다.

그리드는 그림 15-(a)와 같이 열은 가로방향으로 한 단위씩 증가하며 행은 세로방향으로 한 단위씩 증가하는 특성을 가지고 있으며, 각각의 셀을 색인구조에 있는 각각의 셀의 정보를 확인 후 그룹을 형성한다.

Group 1 = {C2, B2, D2, E2, F2, G2}
Group 2 = {B3, B4, B5, B6, B7, B8}



(a) 확장 Cloaking 영역 설정 전



(b) 확장 Cloaking 영역 설정 후

그림 15. 확장 Cloaking 영역 설정

Group 3 = {H3, H4, H5, H6, H7, H8}
Group 4 = {C8, B8, D8, E8, F8, G8}

각 그룹에 사용자 K값을 구한 후 가장 높은 쪽으로 확장을 하면 G3에 K값이 3이기 때문에 그림 15-(b)와 같이 확장이 된다. 만약 각 그룹의 K값이 동일하다면 각각의 해당하는 그다음 행과 열을 확인 후에 K값이 많은 쪽으로 이동한다.

그림 15-(b)와 같이 확장 Cloaking 영역내의 위치한 K값을 검사하여 그 수가 사용자가 요구하는 K값에서 50%이상 차지한다면 수행 단계4로 이동하여 Dummy K를 생성하고, 50%미만이면 수행단계5를 수행하여 계속적으로 Cloaking 영역을 확장해 나간다.

Ⅳ. 구현 및 성능 평가

본 절에서는 L-diversity와 K-anonymity를 고려한 기법 중 Privacy Grid, GBC 그리고 본 논문에서 제안하는 Dummy를 이용한 Cloaking기법(이하 DUCS)을 실험 평가하였고, 실험평가항목은 L-diversity를 만족하는 Cloaking 영역 크기, K-anonymity를 만족하는 Cloaking 영역 크기, 전체 Cloaking 영역 설정 시간으로 각 항목을 비교 평가하였다.

본 논문의 실제 구현 환경은 표 2와 같으며, 구현평가에 사용된 임의의 데이터 생성은 GSTD(Generate Spatio Temporal Data) 알고리즘[10]을 이용하여 2,000개의 건물 L값과 50,000명의 K값을 생성하였고, GSTD 알고리즘은 가로 및 세로의 범위가 1인 사각형 내에 객체를 생성하기 때문에 그리드 크기는 가로 및 세로 크기를 0.01로 고정하여 성능평가를 수행하였다. 그리고 K값과 L값의 배치도에 따라 데이터 값이 각기 다르게 나오기 때문에 여러 번을 수행한 후 결과 값을 평균치로 성능 평가하였다.

표 2. 실험환경

항 목	성 능
CPU	Intel(R) Core(TM)2 Duo CPU 2,40GHz
Memory	4GB
OS	Windows Vista Home Premium K
Compiler	Microsoft Visual Studio.NET 2003

4.1 Cloaking 영역 크기 성능평가

Cloaking 영역 크기 성능평가는 K값을 20이상으로 L값을 4, 6, 8, 10, 12로 증가하여 L값에 따른 변화를

성능 평가 하였고, L값을 8이상으로 K값을 10, 20, 30, 40, 50 씩 증가하여 성능 평가 하였다.

그림 16은 각각의 연구의 Cloaking 영역 크기를 비교한 것으로, DUCS, Privacy Grid, GBC는 모두 L값이 커질수록 Cloaking 영역이 증가하는 것을 확인할 수 있다. 그런데, DUCS가 GBC보다 Cloaking 영역 크기 면에서 더 큰 이유는 모서리로 치우치는 경우, Cloaking 영역이 확장되도록 설계되었기 때문에 GBC보단 영역크기가 큰 것을 확인 할 수 있다.

그림 17은 L값을 8로 하고 K-anonymity값에 따른 변화에 대한 Cloaking 영역 크기를 비교한 데이터 이다. DUCS기법은 Dummy를 이용하기 때문에 Cloaking 영역이 근소하게 변화하는 것을 확인할 수 있지만 Privacy Grid나 GBC 기법은 K값이 증가할수록 더욱 더 영역 크기가 커지는 것을 확인할 수 있다.

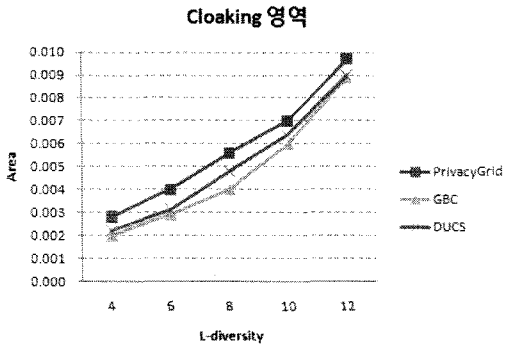


그림 16. L-diversity 증가에 따른 Cloaking 영역

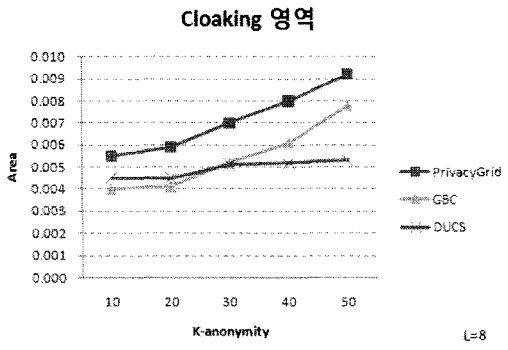


그림 17. K-anonymity 증가에 따른 Cloaking 영역

4.2 Cloaking 영역 설정 시간

본 논문에서 제안한 Dummy를 이용한 Cloaking 영역 설정은 그리드를 사용하여 Cloaking 영역을 설정하기 때문에 빠른 속도를 처리되며, L값이 커질수록 Cloaking 영역 설정 시간이 감소됨을 알 수 있다. 그 이유는 임시 Cloaking 영역에 포함되는 K값이 사용자가 정의한 K' 보다 클 가능성이 높기 때문에 확장 알고리즘을 수행하지 않고 Cloaking 영역 설정이 종료되거나, 본 논문에서 제안한 Dummy K 생성 알고리즘을 수행함으로써 확장 알고리즘을 수행하지 않고 Cloaking 영역 설정이 종료되기 때문이다.

그림 18은 Privacy Grid, GBC, 그리고 DUCS의 Cloaking 영역 설정 시간을 비교한 것이다. 영역 설정 시간 비교는 L값을 8로 고정하고 K값을 10, 20, 30, 40, 50씩 증가시켜 평가하였으며, 그림 18에서 알 수 있듯이 DUCS의 Cloaking 영역 설정시간이 Privacy Grid와의 Cloak빠른 것을 알 수 있다.

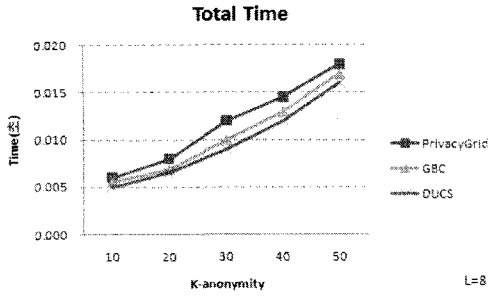


그림 18. Cloaking 영역 설정 시간

V. 결 론

위치기반서비스는 위치정보를 활용해 업무 생산성 개선 및 다양한 생활편의를 제공하는 서비스로 개인 모바일 정보기기인 스마트폰을 활용해 위치정보에 이용자정보, 증강현실 등을 결합함으로써 서비스가 고도화되고 있다. 그러나 사용자의 정확한 위치정보를 DB 서버에 보내기 때문에, 사용자 위치정보가 상대방에서 노출될 수 있다.

본 논문에서는 위치기반 서비스에서 개인의 위치정보를 보호하기 위해 기존에 연구된 L-diversity와 K-anonymity를 만족하는 Cloaking 영역 설정 방법인 Privacy Grid Cloaking 기법[8,9]과 Grid based Cloaking(GBC) 기법[2]을 살펴보고, 각 기법의 문제점을 분석하였다. 그리고 Dummy를 이용한 Cloaking 영역 설정 알고리즘을 설계로 기존 기법의 문제점을 다음과 같이 개선하였다.

첫째, Privacy Grid의 색인구조에 Static Group 항목을 추가시켜 Privacy Grid의 여러 셀에 걸쳐져 있어서 건물을 중복으로 카운트하던 건물 중복 문제를 해결하였다.

둘째, GBC에서 각 건물의 모서리 지점으로 치우쳐서 Cloaking 영역이 설정되는 경우, 각 모서리의 인접한 셀을 검색한 후 K값을 증가시킴으로써 개인 정보 보호를 한층 더 강화시켰다.

마지막으로 사용자가 요구한 K′보다 작은 수의 K가 Cloaking 영역 내에 존재할 때, 영역을 확장하기 전에 Dummy K를 생성하도록 하여 Cloaking 영역 확장시간과 K값을 검색시간을 단축하고, 사용자의 위치 정보 보호를 Dummy를 이용해 한층 더 강화시켰다.

참 고 문 헌

[1] 이성호, "스마트폰과 위치기반서비스를 활용한 서비스산업 혁신전략", *삼성경제연구소 SERI경영노트*, 제62호, 2010.

[2] 엄정호, 김지희, 장재우, "위치기반 서비스에서 개인 정보 보호를 위한 그리드를 위한 Cloaking 영역 생성 알고리즘", *한국공간정보시스템학회 논문지*, 제11권 제2호, pp.151-161, 2009.

[3] M. F. Mokbel, C. Chow, and W. Aref, "The New Casper : Query Processing for Location Services without Compromising Privacy," *In Proc. of the International Conference on Very Large Data Bases*, pp.763-774, 2006

[4] G. Ghinita, P. Kalnis and S. Skiadopoulos, "PRIVE : Anonymous Location-Based Queries in Distributed Mobile Systems," *In Proc. of World Wide Web*, pp.237-246, 2007

[5] L. Sweeney, "K-anonymity : A model for protecting privacy", *International Journal on Uncertainty, Fuzziness and Knowlege-based System*, 10(3), pp.557-570, 2002

[6] A. Machanavajjhala, J.Gehrke, D. Kifer, and M. Venkitasubramaniam, "L-diversity : Privacy beyond k-anonymity", *ACM Transactions on Knowledge Discovery form Data*, 1(1), Article 3, pp.1-52, 2007

[7] Bugra Gedik, Ling Liu, "Protecting Location Privacy with Personalized k-Anonymity : Architecture and Algorithm", *IEEE Transactions on Mobile Computer*, 7(1), pp.1-18, 2008

[8] B. Bamba and L. Liu, "PRIVACY GRID : Supporting Anonymous Location Queries in Mobile Environments", *Research report in National Technical Information Service*, 2007.

[9] B. Bamba and L. Liu, P. Pesti and T. Wang, "Supporting Anonymous Location Queries in Mobile Environments with PrivacyGrid", *The International World Wide Web Conference Committee(IW3C2)*, pp.237-246, 2008.

[10] Yannis Theodoridis, Jefferson R. O. Silva, and Mario A. Nascimento, "On the Generation of Spatiotemporal Datasets," *A TimeCenter Technical Report*, 1999.

김 주 용 (Ju-Yung Kim) 정회원

2007년 2월 관동대학교 컴퓨터공학과 졸업

2009년 2월 관동대학교 전자계산공학과 석사

2009년 3월~현재 관동대학교 전자계산공학과 과정

<관심분야> 모바일 프로그램, 네트워크 보안

정 은 희 (Eun-Hee Jeong) 정회원

1991년 2월 강릉대학교 통계학과 졸업

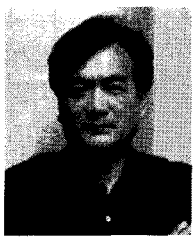
1998년 2월 관동대학교 전자계산공학과 석사

2003년 2월 관동대학교 전자계산공학과 박사

2003년 9월~현재 강원대학교 삼척캠퍼스 지역경제학과 부교수

<관심분야> 네트워크 보안, 전자상거래, 웹 프로그래밍

이 병 관 (Byung-Kwan Lee) 정회원



1975년 2월 부산대학교 기계설계학과 졸업

1986년 2월 중앙대학교 전자계산공학과 석사

1990년 2월 중앙대학교 전자계산공학과 박사

1988년 3월~현재 관동대학교 컴퓨터학과 교수

<관심분야> 네트워크 보안, 컴퓨터 네트워크, 전자상거래