연속적인 위치기반 서비스를 지원하는 분산 그리드 기반 Cloaking 영역 설정 기법 설계

이아름, 김형일, 장재우
전북대학교 전자정보공학부 컴퓨터공학과
e-mail: arlee@dblab.chonbuk.ac.kr, {jwchang,melipion}@jbnu.ac.kr

Distributed Grid-based Cloaking Area Creation Scheme supporting Continuous Location-Based Services

Ah-reum Lee, Hyeong-il Kim, Jae-Woo Chang
Department of Computer Engineering, Chonbuk National University of Korea

요 약

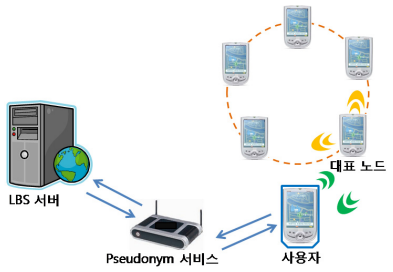
모바일 기기 및 무선 통신 기술의 발달로 인하여 위치 기반 서비스의 이용이 확산되었다. 그러나 이와 같이 사용자의 정확한 위치정보를 가지고 LBS 서버에 서비스를 요청하는 것은 심각한 개인 정보 누출의 위협이 될 수 있다. 따라서 안전하고 편리한 위치기반 서비스 사용을 위한 개인 정보 보호 방법이 요구된다. 이를 위해 본 논문에서는 연속적인 위치기반 서비스를 지원하는 분산 그리드 기반 Cloaking 영역 설정 기법을 설계한다. 설계하는 기법은 분산 환경에서 연속적인 서비스를 지원하기 위해 Cloaking 영역 설정 시 필요한 정보를 분산 유지하고, 이동 확률 매트릭스 생성 및 확률 계산을 분산적으로 수행한다. 마지막으로는 모바일 사용자 사이에 발생하는 통신비용을 감소시키기 위해, 대표 노드는 해당 클러스터에서 떠난 사용자에 대한 정보를 유지하고 클러스터 내 부분 확률값의 합산시 병합노드를 사용한다.

1. 서론

모바일 기기 및 무선 통신 기술의 발달로 인하여 사용자의 위치 정보를 활용한 위치 기반 서비스(Location-Based Service)가 주목받고 있다. 그러나 이러한 서비스는 서비스를 요청하는 사용자가 자신의 정확한 위치정보를 LBS 서버에 보내기 때문에 사용자의 심각한 개인 정보 누출의 위협이 될 수 있다. 따라서 사용자의 안전하고 편리한 위치기반 서비스 사용을 위한 개인 정보 보호 방법이 요구된다. 위치기반 서비스에서 사용자의 위치 정보를 보호하기 위한 대부분의 연구는 K-anonymity를 만족하면서 최소 크기를 가지는 Cloaking 영역을 생성하는 연구를 제안하고 있다. 여기서 K-anonymity(이하 k)는 Cloaking 영역에 질의 요청한 사용자와 그 사용자를 제외한 K-1명의 다른 사용자의 위치 정보를 포함하는 것으로, 사용자의 위치 노출 확률을 1/K로 감소시킨다. 이러한 K-anonymity를 고려한 Cloaking 영역 설정 기법은 모바일 사용자가 이동하면서 연속적으로 서비스를 요청하는 경우, 생성되는 모든 Cloaking 영역에 질의를 요청한 사용자가 존재하지만 다른 사용자는 이동 경로에 따라 Cloaking 영역에 포함되지 않는 경우가 존재한다. 이러한 문제점을 해결하기 위한 기존의 Cloaking 기법 중 가장 우수한 성능을 보이는 연구는 A. Lee et al.이 제안한 GCCA[1]이다. GCCA는 연속적으로 서비스를 요청하는 사용자의 위치 정보를 보호하면서 최소의 Cloaking 영역을 설정하기 위해 Cloaking 영역의 프라이버시 보호 수준인 엔트로피를 계산한 후, 이를 고려하여 Cloaking 영역을 설정한다. 그러나 이 연구는 중앙 집중 방식이기 때문에 anonymizer에서의 병목현상 및 보안 위협 문제가 순간 질의를 처리하는 것에 비해 악화되어 나타날 수 있다. 따라서 본 논문에서는 연속적인 위치기반 서비스를 지원하는 분산 그리드 기반 Cloaking 영역 설정 기법을 설계한다. 향후 연구로는 설계한 분산 기법을 성능평가를 통해 설계한 기법이 기존 연구에 비해 우수함을 제시하고자 한다.

2. 연속적인 위치기반 서비스를 지원하는 분산 그리드 기반 Cloaking 영역 설정 알고리즘

연속적인 위치기반 서비스를 제공하는 기존 연구들 중 가장 우수한 성능을 보이는 연구는 GCCA 이다. 하지만, GCCA 는 중앙 집중 방식을 사용한 연구로써 현재 분산 방식에서 연속적인 서비스를 제공하는 연구는 존재하지 않는다. 따라서 본 논문에서는 기존 연구인 GCCA를 분산 환경으로 확장하여 설계하고자 한다. 분산 환경에서 연속적인 서비스를 제공하기 위해서는 다음과 같은 사항을 고려해야 한다. 첫째, 모바일 기기는 저장할 수 있는 정보가 한정적이므로 연속적인 서비스를 제공하기 위해 필요한 정보들을 분산적으로 유지한다. 둘째, 모바일 기기의 처리 능력 역시 제한적이므로, Cloaking 영역의 보호 수준을 계산하기 위해 사용되는 이동확률 매트릭스 생성과 임시 Cloaking 영역에 포함된 사용자의 확률 계산을 분산적으로 처리한다. 마지막으로는 질의를 요청한 사용자가 이웃한 사용자와의 통신을 통해 Cloaking 영역을 설정하기 때문에 모바일 사용자사이에 발생하는 통신비용이 증가한다. 이를 위해 각 클러스터의 대표노드에서 이전 시간의 떠난 사용자에 대한 정보를 유지하여 이전 사용자를 찾기 위한 통신량을 줄인다. 아울러 클러스터 내 부분 확률값의 합산시 병합노드를 사용함으로써 사용자 사이에 발생하는 통신량을 분배하여 통신비용을 감소시킨다. 본 연구에서 제안하는 시스템 구조는 그림 1과 같다.



(그림 1) 시스템 구조

시스템은 크게 LBS 서버, Pseudonym 서비스, Chord 프로토콜[2]로 구성된다. Pseudonym 서비스는 Cloaking 영역을 LBS 서버에게 보내기 전에 사용자의 IP 주소를 숨겨주는 역할을 하며, LBS 서버는 전송된 Cloaking 영역에 대하여 질의를 처리한다. Chord 프로토콜은 모바일 사용자들이 Chord라는 분산 해시 테이블 구조에 기반을 두어 구성된 가상 링 형태의 P2P 네트워크이며, 링 내의 사용자들은 클러스터 단위로 이루어진다. 본 연구에서는 사용자의 위치를 힐버트 커브를 이용하여 힐버트 값으로 변환하고 이를 인덱싱 키로 고려하여 링 구조를 형성하는 클러스터 구성방법[3]을 사용한다. 아울러 연속적인 서비스를 지원하기 위해 일정 시간 후 모든 사용자들의 위치 이동을 고려하여 클러스터를 재구성한다고 가정한다. 이때, 클러스터의 대표노드

는 사용자사이의 발생하는 통신비용을 줄이기 위해 이전 시간에 클러스터를 떠난 사용자의 정보를 유지한다. 이를 통해 사용자가 질의 요청자를 검색할 경우, 대표노드가 유지하고 있는 정보를 통해 사용자를 효율적으로 찾을 수 있다.

본 연구에서는 GCCA를 분산 환경으로 확장한 연구를 수행한다. 이를 위해, A. Lee et al.의 연구인 DAHC[4]를 이용하여 힐버트 값으로 암호화된 사용자의 위치정보를 계산하여 인접한 셀을 존재하는 이웃한 사용자를 검색하는 방법을 사용한다. 설계하고자 하는 Cloaking 기법의 수행단계는 다음과 같다.

수행단계 1. 임시 Cloaking 영역 설정

서비스 요청 시(T=0), DAHC를 사용하여 Cloaking 영역이 설정되었다고 가정할 때, 일정 시간이 지나면 설정된 Cloaking 영역 안에 존재하는 사용자들은 각자의 이동 경로를 따라 움직인다. 이후, 갱신된 위치정보를 질의 요청자에게 전송한다. 질의 요청자는 전송받은 정보를 바탕으로 임시 Cloaking 영역을 설정한다. 임시 Cloaking 영역이 설정되면, 질의 요청자는 임시 Cloaking 영역에 속한 모든 사용자를 찾기 위하여 자신을 중심으로 이웃한 셀부터 임시 Cloaking 영역에 속한 모든 셀을 계산한다. 계산된 이웃셀 정보를 클러스터의 대표노드에 전송하여 이웃한 사용자가 존재하는 지를 검색하여 정보를 얻어온 후, 이를 기반으로 사용자 정보 테이블을 생성한다. 사용자 정보 테이블은 프라이버시 보호 수준을 고려한 Cloaking 영역 설정 시, 가까운 사용자들을 우선적으로 고려하기 위한 것으로, 임시 Cloaking 영역에 속한 사용자의 ID와 위치정보인 힐버트 값을 질의 요청자로부터 가까운 순으로 저장한다.

수행단계 2. 임시 Cloaking 영역에 존재하는 사용자의 확률값 계산

설정되는 Cloaking 영역의 프라이버시 보호 수준인 엔트로피를 계산하기 위해 이전 시간에 설정된 Cloaking 영역에 속한 사용자(이후 참여자)들이 질의 요청자일 확률값을 계산한다. 사용자들의 확률값을 구하기 위해서는 이동확률 매트릭스(M)를 생성해야 하는데, 이전 Cloaking 영역에 속한 사용자의 정보를 반영하기 위하여 사용자가 Cloaking 영역에 속했던 횟수를 고려한다. 이를 기반으로 이전 시간의 어떤 사용자가 나중 시간의 어떤 사용자의 위치로 움직일 수 있는 α개의 샘플을 생성하고, 이를 이동확률 매트릭스 생성 시 반영한다. 사용자들의 확률값을 분산적으로 계산하기 위하여 참여자들에게 임시 Cloaking 영역에 속한 사용자들의 ID 목록과 병합노드, 병합할 사용자 수, α에 대한 정보를 전송한다. 병합노드란, 해당 클러스터에 속한 참여자의 부분 확률값을 합산하는 노드를 말한다. 병합노드는 질의 요청자가 참여자들에게 정보 전송 시 선정되며 병합할 사용자의 수만큼 통신을 기다린 후 전송된 정보를 합산하여 질의 요청자에게 전송한다. 참여자들은 질의 요청자로부터 전송받은 정보를 기반으로 자신이 유지하고 있는 count 정보와 확률 값을 가지고 부분 이동확률 매트릭스를 생성한다. 이후, 식(1)를 이용하여 부분 확률값을 구한다.

$$U_{prob} = \sum_{i=1}^{m} (before\, U_{prob} * \frac{M_{ij}}{\alpha}) \cdots\cdots\cdots\cdots\cdots\cdots (1)$$

여기에서 U_{prob}는 부분 확률값을 의미하고, beforeU_{prob}는 이전 시간에서의 확률값을 의미한다. 참여자들은 계산된 부분 확률값을 병합노드로 전송하고, 각 클러스터내의 병합노드에서는 부분 확률값을 합산하여 질의 요청자에게 전송한다. 질의 요청자는 각 클러스터의 병합노드로부터 전송받은 부분 확률값을 합산하여 임시 Cloaking 영역에 포함된 사용자의 확률값을 구한다. 참여자들의 부분 확률값을 각 클러스터의 병합노드에서 합산하여 질의 요청자에게 전송하는 것이 참여자들이 질의 요청자에게 직접 전송하는 것에 비해 발생하는 통신비용을 감소시킬 수 있다.

수행단계 3. 프라이버시 보호 수준을 고려한 Cloaking 영역 설정

임시 Cloaking 영역에 속한 사용자들의 확률값이 계산되면, 이를 고려하여 프라이버시 보호 수준인 엔트로피를 계산한다. 임시 Cloaking 영역 설정 시, 생성한 사용자 정보 테이블을 기반으로 참여자의 수가 질의 요청자가 요구하는 k와 같을 경우, 포함된 사용자들의 확률값을 기반으로 엔트로피를 계산한다. 계산된 엔트로피가 질의 요청자가 요구한 k값보다 클 경우, 사용자들을 포함하는 최소 경계 사각형을 Cloaking 영역으로 설정한다. 만약, k값을 만족하지 않는다면 k 이상의 사용자를 고려하여 엔트로피를 계산한다. Cloaking 영역이 설정되면, 각 참여자에게 계산된 확률값을 전송하여 참여자들이 유지하고 있는 확률값과 count정보를 갱신하게 한다. 이는 이후 Cloaking 영역을 설정하는데 이용한다. 수행단계를 고려한 알고리즘은 분산적으로 프라이버시 수준을 고려하여 Cloaking 영역을 설정하기 때문에 질의 요청자와 참여자 측면으로 구분하여 그림 2와 같이 설계한다.

```
//입력: 질의 요청자 좌표정보(q), K-anonymity(k), 서비스 수행시간(T), T=0인 경우 Cloaking 영역
//출력: 각 T에 해당하는 k를 만족하는 Cloaking 영역
질의 요청자 측면
1. For(t(현재 서비스 시간) < T)
2. {
3.    참여자들로부터 갱신된 위치정보를 받아 임시 Cloaking영역 설정
4.    For(임시 Cloaking 영역에 속한 모든 셀)
5.    {
6.       q의 셀을 중심으로 이웃한 셀을 계산함
7.       계산된 셀 정보를 기반으로 이웃한 사용자 검색
8.       If(사용자가 존재) then 사용자 정보 테이블에 정보를 추가
9.    }
10.   참여자들에게 부분 확률값을 계산하기 위한 정보를 전송함
11.   For(참여자들의 수)   //참여자들의 통신을 기다림
12.   {
         전송받은 값을 현재 계산된 임시 Cloaking 영역에 존재하는
         사용자들의 확률값에 더함
13.   }
14.   사용자 정보 테이블을 기반으로 q를 포함한 k명의 사용자를 선택
15.   Do
16.   {
17.      선택된 사용자를 고려하여 프라이버시 보호 수준(D)을 계산
18.       If(D>= k){
20.         선택된 사용자를 포함하는 최종 Cloaking 영역 설정
21.         설정된 Cloaking 영역의 사용자들에게 계산된 확률값 전송
22.       }
23.      사용자 정보 테이블에 선택된 사용자의 수를 증가함
24.   }While(D<K)
25. }
참여자 측면
//유지: 이전 시간의 사용자 확률값, Count정보
26.  질의 요청자로부터 부분 확률값을 계산하기 위한 정보를 받음
27.  Count정보와 전송받은 정보인 임시 Cloaking 영역에 속한
     사용자들의 ID목록을 기반으로 부분 이동확률 매트릭스를 생성
28.  부분 이동확률 매트릭스를 고려하여 부분 확률값을 계산
29.  If(참여자 == 병합노드)
31.    병합할 사용자의 수만큼 통신을 기다려 부분 확률값을 합산
32.    질의 요청자에게 합산된 결과를 전송함
34.  Else
35.   병합노드에게 계산된 부분 확률값을 전송함
```

(그림 2)연속적인 위치기반 서비스를 지원하는 분산 그리드 기반 Cloaking 영역 설정 알고리즘

Acknowledgments

이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(No. 2009-0059417)

참 고 문 헌

[1] 이아름, 김형일, 장재우, "연속적인 위치기반 서비스를 지원하는 그리드 기반 Cloaking 영역 설정 기법", 한국공간정보시스템 학회 제11권, 3호, 2009.10

[2] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek and H. Balakrishnan, "Chord: A Scalable Peer-to-peer Lookup Service for Internet Application," In Proc. of IEEE/ACM TON, Vol.11 No.1, 2003, pp. 17-32.

[3] G. Ghinita, P. Kalnis and S. Skiadopoulos, "MobiHide: A Mobilea Peer-to-Peer System for Anonymous Location-Based Queries," In Proc. of SSTD, Vol.4605, 2007, pp. 221-238.

[4] 이아름, 엄정호, 장재우, "분산 그리드 환경에서 힐버트 커브를 이용한 효율적인 Cloaking 영역 설정기법", 한국공간정보시스템 학회 제11권, 1호, 2009.03