
의미 공간에서의 실내 측위를 위한 베이지안 알고리즘

김희겸* 탁성우*

*부산대학교

Bayesian Algorithm for Indoor Semantic Location Determination

Hee-Kyum Kim* · Sung-Woo Tak*

*Busan National University

E-mail : shlrur123@gmail.com

요 약

최근 스마트 폰과 같은 무선 모바일 기기 사용량의 증가 때문에 위치 기반 서비스(LBS : Location Based Service)에 대한 연구가 활발히 증가하고 있다. 실외 위치 측위는 무선 모바일 기기에 내장된 GPS를 이용할 수 있다. 하지만 일반 건물보다 규모가 큰 대형 크루즈선의 실내 같은 곳에서는 GPS 사용이 불가능하므로 실내 환경에 적합한 위치 측위 방식을 고려하여야 한다. Wi-Fi(Wireless Fidelity)는 실내의 여러 곳에 설치 되어 있어서 추가로 Wi-Fi의 설치과정을 거치지 않아도 되고, 다른 무선 센서 기기와 비교하면 비교적 싼 가격을 가진다. 본 논문에서는 Wi-Fi의 신호를 이용하여 실내에서 의미 있는 공간을 인지하는 베이지안 알고리즘을 소개한다.

ABSTRACT

As the amount of the wireless mobile products like a 'Smart phone' used increases, the studies about the Location Based Service (LBS) is highly increasing. Outdoor location determination can use the GPS which is built-in in the wireless mobile products. However, it is not possible to use GPS inside the huge cruise bigger than a normal building, it is regarded to consider Indoor location determination which is appropriate at the inside environment. Wi-Fi (Wireless Fidelity) does not need an extra installation process because it is already installed here and there inside the building. In this respect, Wi-Fi has low price competitiveness compared to other wireless sensor products. In this paper, I will introduce 'Bayesian Algorithm' which can recognize useful space with Wi-Fi signal.

키워드

Wi-Fi, LBS, 베이지안 알고리즘, 실내 측위, 의미 있는 공간

I. 서 론

최근 위치 기반 서비스(Location Based Service: LBS)는 여러 분야와 결합하여 다양한 서비스가 제공되고 있는데, 이를 위한 기반 기술로서 측위(Localization)의 중요성이 주목받고 있다. 특히, 실내 지역은 GSP(Global Positioning System)[1] 신호의 수신에 불가능하여 무선 센서 네트워크를 활용한 측위 알고리즘이 활발하게 연구되고 있다.[5]

실내 측위 방법으로는 RFID[2], 초음파[3], UWB[4]등을 이용한 방법이 있다. 하지만 이러한 위치 측위 기술들은 대부분 추가로 많은 장비 및

설치비용이 필요한 단점을 가지고 있다.

이에 반해서 최근 새로운 실내 위치 측위 방식으로 Wi-Fi(Wireless Fidelity)를 사용하는 방법이 대두되고 있다. 무선랜은 현재 가장 널리 쓰이고 있는 무선 통신 방식 중 하나로 이미 여러 분야에서 안정성이 검증되어 있다. 곳곳에 설치되어 있는 무선 인터넷 망을 활용하여 실내 위치를 식별하기 때문에 적은 비용으로도 시스템을 구축할 수 있으며 외부 인터넷 연결이 용이한 AP(Access Point)를 활용하기 때문에 위치 측정과 동시에 인터넷 통신도 사용할 수 있다.

현대에는 크루즈나 항공모함 등 대형 선박이

존재한다. 특히 대형 크루즈 선박은 그 규모가 백화점과 같은 대형 건물과 비교해서 작다고 할 수 없다. 로얄캐리비안 사의 오아시스 호를 예로 보면, 길이/너비가 362m/47m 이고 실내는 총 18층에 이른다. 이처럼 대형 크루즈 선박의 실내 환경에서는 실내 측위 기술이 반드시 필요하며, 크루즈 선박 내에서 승객들의 인터넷 통신을 위해 쓰이는 무선 AP를 사용한 측위 기술이 비용과 효율적인 면에서 알맞다고 볼 수 있다.

Wi-fi를 이용하는 여러 가지 방법 중 fingerprinting 방식을 채용하는 것이 가장 효율적인 것으로 알려져 있다.[6] Fingerprinting 방식은 위치 측위 대상이 되는 공간을 일정한 범위로 나누고 각 RP(Reference Point)의 위치 값을 데이터베이스에 저장한 후 위치 값과 함께 AP로부터 단말기(MU: Mobile Unit)에 도달한 신호세기를 측정하여 대상이 되는 공간의 모든 RP를 측정할 때까지 반복적으로 수행된다.

본 논문에서는 fingerprinting 방식의 측위 대상이 되는 공간을 일정한 범위로 나누는 대신 의미 공간이라는 개념을 사용해서 실내 위치 측위를 하였다. 단말기에 도달한 AP의 RSSI 신호를 bayesian 알고리즘을 사용하여 단말기가 있는 의미 공간을 알아낸다. 그리고 bayesian 알고리즘을 사용하여 위치 측위를 하는 시스템의 학습을 통해서 결과에 대한 신뢰도를 향상시킬 수 있다.

본 논문에서는, 단말이 있는 의미 공간을 알아내는 bayesian 알고리즘을 제안하고, 알고리즘의 신뢰도를 향상시키는 학습 방법을 제안한다.

II. 의미 공간

실외는 대부분 장애물이 없는 열린 공간이라는 것을 특징으로 할 수 있다. 이에 비해서 실내는 벽과 같은 많은 장애물이 있는 환경이기 때문에 path loss가 많게 되고 위치 측위를 하는데 실외에서의 위치 측위에 비해서 더 많은 오차가 포함된 결과가 나올 수 있다. 하지만 실내에서 벽이 없는 공간에서의 path loss는 그리 크지 않게 된다.

실내 위치 공간은 실외 위치 공간과는 다르게 그 공간을 사용하는 사용자에게 의해서 벽으로 나뉘어져 있다. 그림 1은 부산대학교 자연대 연구실 동 3층의 일부분을 그림으로 나타낸 것이다. 여기서는 사용 목적을 위해서 1, 2, 5, 8지역의 복도와 3지역의 교수 연구실, 4지역의 연구실1, 6지역의 unlab 연구실, 7지역의 전산실, 9지역의 연구실2가 벽으로 나누어져 있는 것을 알 수 있다. 이렇게 사용자들이 목적을 가지고 의미 있게 나누어져 있는 공간을 의미 공간이라고 한다.

실내 위치는 이미 벽으로 나누어진 공간이기 때문에 실외의 열린 공간에서처럼 x, y, z좌표로 측위 할 필요가 없다. 하나의 의미 공간의 크기가 사용자의 인식이 힘들 정도로 크지 잘 나누어져

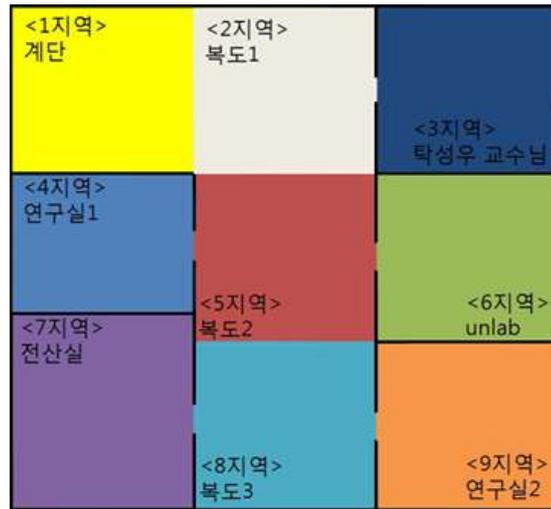


그림 1. 의미 공간

있다면, 실내 위치 측위의 결과가 하나의 의미 공간이 될 수 있다.

III. Bayesian Decision Algorithm

Bayesian 알고리즘은 실제(Evidence) 확률과 추정(Hypothesis) 확률, 그리고 실제일 때 추정하였을 확률을 바탕으로 추정하였을 때 실제일 확률을 구하는 것이다. 이를 식으로 표현한 것이 식(1)이다.

$$P(E|H) = \frac{P(H|E) \cdot P(H)}{P(E)} \quad \text{수식(1)}$$

수식(1)에서 $P(E|H)$ 는 추정하였을 때 실제일 확률을 뜻하고 $P(H|E)$ 는 실제일 때 추정하였을 확률을 뜻한다. $P(H)$ 는 추정할 확률을 뜻하고 $P(E)$ 는 실제일 확률을 뜻한다.

Bayesian 알고리즘은 실내 위치 측위에서는 수식(2)와 같이 사용될 수 있다.

$$\begin{aligned} &P(n\text{지역 실제} | m\text{지역 추정}) \quad \text{수식(2)} \\ &= \frac{P(m\text{지역 추정} | n\text{지역 실제})}{P(m\text{지역 추정})} \\ &\quad \times P(n\text{지역 실제}) \end{aligned}$$

수식(2)에서 $P(n\text{지역 실제} | m\text{지역 추정})$ 은 m 지역으로 추정 하였을 때 실제로는 n지역에 있을 확률을 뜻한다.

수식(2)의 $P(m\text{지역 추정} | n\text{지역 실제})$ 는 n지역에 사용자가 있을 때, m지역에 있다고 추정하는 것을 뜻한다. 이 때는 의미공간에 따라서 저장되어있는 AP들의 RSSI 값과 현재 탐색된 AP들의 RSSI 값들을 유클리디안 거리방식으로 가장 추정이 높게 나온 지역의 표준편차 S_i 를 사용하여 구하게 된다. 총 의미공간이 n개라고 할 때, $P(m\text{지역 추정} | n\text{지역 실제})$ 는 수식(3)을 통하여 구하게 된다.

$$P(m\text{지역 추정} | n\text{지역 실제}) \quad \text{수식(3)}$$

$$= \frac{S_i}{\sum_{x=1}^n S_x \times \sum_{x=1}^n S_x^{-1}}$$

수식(2)의 $P(n\text{지역 실제})$ 는 n 지역에 실제 있을 확률을 뜻한다. 이 확률은 확률적 계산을 통해서 구할 수 없는 값이다. 실제 사용자가 의미 공간인 n 지역에 실제로 있을 확률을 나타내는 값이기 때문에 n 지역에 사용자가 얼마나 자주 있을 거라는 통계적인 자료가 필요하다. 이러한 자료는 다음 장에서 다루게 되는 bayesian 학습 알고리즘에서 다루는 값을 사용하게 된다. 그림 1을 예로 들면 9개의 의미 공간이 가지는 n 지역에 실제 있을 확률은 아무 정보가 없고 모든 의미 공간이 동등한 확률을 가지기 때문에 $1/9$ 가 된다. 하지만 학습 횟수가 늘어날수록 수식(4)와 같은 값을 가진다.

$$P(n\text{지역 실제}) \quad \text{수식(4)}$$

$$= \frac{l_i + 1}{\sum_{i=1}^n l_i + 9}$$

수식(4)에서 l_i 는 실제 i 지역에서 학습을 한 횟수를 뜻한다.

수식(2)의 $P(m\text{지역 추정})$ 은 m 지역에 있다고 추정하는 확률을 뜻한다. m 지역에 있다고 추정할 수 있는 확률은 실내 위치 공간에 따라서 정해지게 된다. 그림 1을 예로 들면 1지역에 인접한 지역은 2, 4, 5지역이 되기 때문에 자신의 지역을 포함해서 4개의 지역 중에 1지역에 있을 확률로 1지역에 있다고 추정할 수 있는 확률은 $1/4$ 을 가지게 된다. 마찬가지로 5지역은 다른 모든 지역과 인접해 있으므로 5지역에 있다고 추정할 수 있는 확률은 $1/9$ 을 가지게 된다.

수식(2)의 결과 값은 의미공간의 개수만큼 얻을 수 있다. $P(n\text{지역 실제} | m\text{지역 추정})$ 에서 m 지역은 유클리디안 거리 방법을 사용해서 구할 수 있다. 그리고 n 지역에 의미공간을 모두 대입하여서 $P(n\text{지역 실제} | m\text{지역 추정})$ 값이 가장 높은 확률을 가지는 n 지역을 현재 사용자가 있을 확률이 가장 높은 의미 공간으로 인식하게 된다.

IV. Bayesian Learning Algorithm

Bayesian 알고리즘을 사용하여 위치 측위를 하는 시스템에서는 2개의 값을 가지고 학습을 하게 된다.

$Ref_{i,ssid}$ 는 i 번째 의미공간 AP(ssid를 가지는)의 rssi reference값을 뜻한다.

$cnt_{i,ssid}$ 는 i 번째 의미공간 AP(ssid를 가지는)의 학습 횟수를 뜻한다.

2개의 값은 각 의미 공간에서 학습을 할 때 마

다 갱신된다. 각 값의 갱신은 수식(5)와 수식(6)에 나타나 있다.

$$Ref_{i,ssid} \quad \text{수식(5)}$$

$$= \frac{Ref_{i,ssid} \cdot cnt_{i,ssid} + RSSI_{ssid}}{cnt_{i,ssid} + 1}$$

$$cnt_{i,ssid} = cnt_{i,ssid} + 1 \quad \text{수식(6)}$$

수식(5)에서 $RSSI_{ssid}$ 는 ssid를 가지는 AP의 rssi값을 뜻한다.

$Ref_{i,ssid}$ 에서 알 수 있듯이, 학습을 할 때는 지역 i 와 AP의 ssid값, rssi값을 정확하게 대입해 주어야 큰 오차가 발생하지 않는다.

수식(5)의 학습은 유클리드 거리 방식을 위한 학습이다. Bayesian 알고리즘을 위한 학습 방법은 수식(2)의 $P(n\text{지역 실제})$ 값을 바꾸어 준다.

i 지역에서 학습을 할 때마다 $P(i\text{지역 실제})$ 값과 $P(i\text{를 제외한 나머지 지역 실제})$ 는 수식(7)과 수식(8)과 같이 갱신된다.

$$P(i\text{지역 실제}) = \frac{l_i + 1}{\sum_{x=1}^n l_x + 1} \quad \text{수식(7)}$$

$$P(\text{나머지 지역 실제}) = \frac{l_j}{\sum_{x=1}^n l_x + 1} \quad \text{수식(8)}$$

수식(7)에서 l_i 는 i 지역의 과거 학습 횟수를 뜻한다. 수식(8)의 j 는 i 가 아닌 나머지 지역들을 뜻한다.

V. 결 론

본 논문에서는 실내 위치 측위에서 의미 공간이라는 개념을 사용하고 bayesian 알고리즘의 결정과 학습 알고리즘에 대해서 제안 하였다. bayesian 결정 알고리즘에서 각 확률이 가지는 의미와 계산 방법을 제안 하였다.

의미 공간이라는 개념을 실내 위치 측위에 도입하게 된다면 위치오차가 낮고 사용자가 필요로 하는 위치 값을 제공할 수 있을 것이다.

향후 과제로서 본 논문에서 제안한 알고리즘을 사용하여 실내 측위 실험을 하고 결과 값에 대한 분석을 할 예정이다.

참고문헌

- [1] R. Bajaj, S. L. Ranaveera and D. P. Agrawal, "GPS Location Tracking Technology", IEEE Computer, Vol. 35, pp. 92-94, 2002.
- [2] Wireless Technologies for Ubiquitous Services. NTT Review December 2003, Vol. 1, No. 94.

- [3] A. Harter, A. Hopper, P. Steggles and P. Webster, "The anatomy of context-aware application", Proceeding of fifth annual ACM/IEEE international conference on Mobile computing and networking, 1999.
- [4] S. Gezici, Z. Tian, G. B. Biannakis, H. Kobayashi, A. F. Molish, H. V. Poor and Z. Sahinoglu, "Localization via ultra-wideband redios:a look at positioning aspects for future sensor networks", IEEE signal processing magazine, Vol. 22, pp. 70-84, 2005
- [5] 이장재, 권장우, 정민아, 이성로, "실내 측위 결정을 위한 Fingerprinting Bayesian 알고리즘", 한국통신학회 논문지, Vol. 35, No. 6, pp.888-894, 2010.
- [6] P. Bahl and V.N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System", Proc. IEEE Computer and Communication Societies, Vol. 2, pp. 775-784, 2000.