

논문 2006-43CI-6-1

옥내 측위를 위한 지문 방식 알고리즘들의 성능 분석

(Performance Analysis of Fingerprinting algorithms for Indoor Positioning)

임 재 결*

(Jaegel Yim)

요 약

무선네트워크 기반 옥내측위는 측위를 위한 특수 장비를 필요로 하지 않고, 지문 방식은 무선네트워크 기반 측위를 위한 기술 중에서 가장 정확도가 높기 때문에 무선네트워크 지문 방식이 가장 적당한 옥내측위 방법이다. 지문 방식은 준비 단계와 실시간 측위 단계로 구성되는데, 준비 단계는 그렇지 않지만 실시간 측위 단계는 실행 시간이 매우 중요한 요소인 작업이다. 왜냐하면, 실시간 측위 단계의 실행 시간이 너무 길면, 실행 도중에 사용자가 너무 많이 이동하여 정확한 측위가 불가능하게 되기 때문이다. 그럼에도 불구하고 무선네트워크 기반 지문 방식의 실시간 측위 단계의 효율성을 개선하는 연구는 아직 수행된 바가 없다. 이 논문은 무선네트워크 기반 지문 방식을 위한 판단나무 방법을 제안하며, 기존의 K-NN 방법 그리고 베이지안 방법과 본 논문이 제안하는 판단나무 방법을 정확도와 실행 시간 측면에서 비교 분석한다.

Abstract

For the indoor positioning, wireless fingerprinting is most favorable because fingerprinting is most accurate among the techniques for wireless network based indoor positioning which does not require any special equipments dedicated for positioning. The deployment of a fingerprinting method consists of off-line phase and on-line phase. Off-line phase is not a time critical procedure, but on-line phase is indeed a time-critical procedure. If it is too slow then the user's location can be changed while it is calculating and the positioning method would never be accurate. Even so there is no research of improving efficiency of on-line phase of wireless fingerprinting. This paper proposes a decision-tree method for wireless fingerprinting and performs comparative analysis of the fingerprinting techniques including K-NN, Bayesian and our decision-tree.

Keywords: 옥내측위, 지문방식, K-NN, 베이지안, 위치기반서비스

I. 서 론

사용자의 사용 환경을 고려하여 걸 맞는 서비스를 제공하는 것을 상황인식(context-aware) 서비스라고 한다. 예를 들어, 사용자의 단말기의 크기와 해상도를 인식하여 이미지의 크기와 해상도를 조절하여 제공하는 것도 일종의 상황인식 서비스이다. 상황인식 서비스의 대표적인 형태로 위치기반서비스를 예로 들 수 있다. 위치기반 서비스(LBS: Location Based Service)란 사용자가 현재 위치한 지리적 위치를 고려하여 사용자에게

유용한 정보를 제공하는 서비스를 일컫는다. 위치기반 서비스의 예로 자동차 항법 시스템이 널리 사용되고 있다. 여기에서 지리적 위치란 일반적으로 위도, 경도, 고도로 결정되는 지구상의 좌표를 일컫지만, 실내에 있는 사용자에게 제공되는 서비스의 경우에는 도면상의 X, Y, Z 좌표가 된다. 위치기반서비스의 예로 공공안전 서비스, 위치추적 서비스, 항법 서비스, 정보제공 서비스 등 부가가치가 높은 서비스들이 많이 있는데^[1], 이러한 서비스를 개발 하려면 필수적으로 사용자의 위치를 파악해야 한다. 사용자의 위치를 파악하는 것을 측위(positioning)라고 한다.

측위의 대표적인 수단으로 GPS와 GLONASS가 널리 사용된다^[2]. GPS는 미국이, GLONASS는 러시아가

* 평생회원, 동국대학교 컴퓨터멀티미디어학과
(Dept. of Computer and Multimedia, Dongguk Univ.)

접수일자: 2006년5월17일, 수정완료일: 2006년10월30일

제공하는 인공위성 기반 측위데이터 제공 시스템이다. 측위 위성은 항상 지구를 향하여 전파를 발사하고 있으며, GPS 수신 장치는 이 전파를 참조하여 위성에서 송신한 신호가 수신 장치에 도달하는데 걸리는 시간을 계산하고, 이 시간에 광속을 곱하여 수신 장치까지의 거리를 계산한다. 지구상의 어떠한 수신 장치라도 4 개의 위성으로부터 동시에 신호를 받을 수 있도록 미국 (러시아) 국방성은 24개의 측위 위성을 항상 작동시킨다. 4 위성의 위치와 수신 장치까지의 거리를 바탕으로 삼각 측량법을 적용하여, 수신 장치는 자신의 위치를 계산한다. GPS 수신기의 원리나 제작 방법은 참고문헌^[3-6]을 참조하기 바란다.

위치기반 서비스는 사용자에게 매우 유용한 정보를 제공함으로써 사용자가 건물 내에 있는 밖에 있는 사용자의 위치를 파악하여 위치기반 서비스를 제공하여 주는 것이 바람직하다. 그러나 GPS 수신기는 위성과 수신장치 사이에 아무런 장애물이 없어야만 (Line-of-Sight) 측위 위성 데이터의 수신에 가능하다. 따라서 건물 내부에 있는 사용자의 위치를 파악하는데 GPS 수신기는 거의 아무런 도움을 주지 못한다. 그래서 측위는 옥외 측위와 옥내 측위로 구분되고, 옥외 측위는 GPS 방법을 중심으로 발전하고 있으며, 옥내 측위는 옥외 측위와 별도로 발전하고 있다. 옥내 측위는 옥내 위치기반 서비스 시스템 구축에 열쇠가 되는 기술이다. 옥내 위치기반서비스의 예로 백화점에서 고객을 원하는 상품이 진열된 곳으로 안내하는 서비스, 박물관에서 고객이 위치한 곳에 진열된 전시물에 대한 콘텐츠를 고객에게 능동적으로 출력하여주는 서비스, 등 무수히 많은 예들 수 있다.

다양한 옥내 측위 시스템이 발표되었다. 예를 들면, Active Badge^[7] 시스템에서는 이동물체들에 적외선을 발사하는 송신기가 부착되어 있고, 방마다 여러 개의 수신기가 고정 지점에 설치되어 있다. 적외선은 벽을 통과하지 못하기 때문에 이동물체가 어느 방에 들어오면 그 방에 설치된 고정 수신기들만 그 이동물체의 적외선을 감지할 수 있다. 수신기는 중앙컴퓨터에 연결되어 있어서 수신기에 감지된 적외선을 바탕으로 이동물체가 어느 방 안에 있다는 것을 판정한다. 이상에서 본 바와 같이, Active Badge 시스템에서는 측위를 위한 특수 장비가 필요하고, 측정한 위치의 단위가 방이라는 것을 알 수 있다.

BAT^[8,9] 시스템에서는 이동물체에 부착된 송신기가 고주파 (RF: Radio frequency) 신호와 초음파

(Ultrasound)를 송신한다. 수신기들은 천정에 행렬 모양으로 고정되어 있다. 송신기가 고주파 신호와 초음파 신호를 동시에 송신하면, 수신기는 이 두 가지 신호가 도착한 시각의 차이를 바탕으로 송신기와 수신기 사이의 거리를 구한다. 번개가 보인 시각과 천둥이 들린 시각의 차이로 번개가 친 장소와 나와의 거리를 구할 수 있는 것과 동일한 원리를 적용하는 것이다. 수신기들은 모두 중앙컴퓨터에 연결되어 있어서 중앙컴퓨터는 수신기들의 좌표와 이동물체와의 거리를 바탕으로 이동물체의 정확한 위치를 판정한다. 이상에서 본 바와 같이, BAT 시스템도 Active Badge와 마찬가지로 측위를 위한 특수 장비를 사용함으로써 설치 경비가 많이 든다는 것을 알 수 있다.

Cricket^[10] 시스템에서는 Active Badge나 BAT와 반대로 이동물체 자신이 자신의 위치를 계산한다. Cricket도 BAT와 마찬가지로 고주파 신호와 초음파 신호를 함께 사용하는데, BAT에서는 이동물체에 부착된 송신기가 신호들을 송신한 반면, Cricket에서는 고정 지점에 부착된 비컨 (beacon)들이 고주파 신호와 초음파 신호들을 동시에 송신하면 이동물체에 부착된 수신기가 이 신호들을 수신하고, 신호들의 도착 시각의 차이를 바탕으로 거리들을 구하고, 비컨들의 좌표와 거리를 바탕으로 자신의 좌표를 구한다. 이상에서 본 바와 같이, Cricket 시스템도 역시 비컨과 수신기와 같은 측위를 위한 특수 장비를 사용한다.

RADAR^[11] 시스템에서는 이동물체가 고정 지점에 위치한 세 개의 기지 스테이션 (base station: 탁상 컴퓨터를 사용함)에 UDP 패킷을 송신하면 기지 스테이션이 패킷의 신호 세기를 바탕으로 이동물체의 위치를 판독한다. 판독을 위하여 사용되는 기술은 K-NN (K nearest neighbor)이라는 일종의 지문(fingerprint) 방식이다. 지문 방식은 준비단계 (Off-line phase)와 실시간 측정 단계 (On-line phase)로 구성된다. 준비단계에서는 측위 소프트웨어가 적용되는 범위내의 각 지점(후보지점이라고 함)에서 특징 값(RADAR에서는 UDP packet의 신호의 세기)을 측정하여, 측정치들의 평균으로 찾아보기 표를 작성한다. 실시간 측정 단계에서는 사용자의 현재 위치에서 측정한 특징 값과 찾아보기 표의 내용과 비교하여 사용자의 위치를 판정한다. 이때 K 개의 가장 가까운 지점을 찾아, 이들의 평균을 사용자의 현재 위치로 판정하는 방법을 K-NN이라 한다.

RADAR 방법을 예를 들어 설명하면 다음과 같다. 옥내 측위 대상 지역의 건물 배치도가 그림 1과 같다고

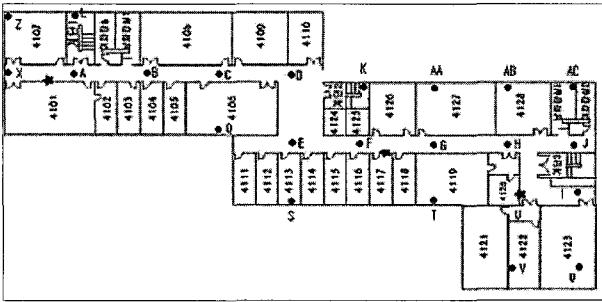


그림 1. 옥내 측위 대상 지역인 건물 4층의 배치도
 Fig. 1. Configuration of the 4th floor, the test bed of our experiment.

표 1. 찾아보기 표의 예
 Table 1. An example Lookup Table.

후보지점	RSS1	RSS2	RSS3
A	31	48	55
B	37	47	54
C	40	45	52
D	45	45	50
...

하자. 기지 스테이션 세 개를 그림의 별표처럼 거리를 두어 임의로 배치한다. 그림의 까만 점(Z, X, A, ... 등의 label이 붙어 있음)을 후보지점이라 하자. RADAR에서는 실시간 측정 단계에서 표 1과 같이 각 후보지점에서 측정된 특징 값으로 구성된 찾아보기 표를 작성한다. RADAR에서는 수신 신호의 세기(RSS: Received Signal Strength)가 특징 값으로 사용된다. 표에서 RSS1은 첫 번째 기지 스테이션에서 측정된 신호의 세기이다.

RADAR에서는 실시간 측위 단계에서 특징 값을 실측하여 찾아보기 표에서 가장 가까운 항을 K 개 찾는다. 예를 들어 실시간 측위 단계에서 실측한 값이 X=(38, 46, 53)이라 하자. X와 A의 거리는 $\sqrt{(38-31)^2 + (46-48)^2 + (53-55)^2}$ 이 된다. 비슷한 방법으로 B, C, D 각각에 대한 거리를 구하면, B가 가장 가깝고, 다음은 C가 두 번째로 가깝다는 것을 알 수 있다. K가 1이라면 X와 가장 가까운 점 하나를 찾아 현재 위치가 B라고 판정한다. K=2라면 가장 가까운 두 점, B와 C를 찾아 이들의 평균 (지점이 좌표임으로 좌표의 평균)을 현재 지점으로 판정한다. 이상에서 살펴 본 바와 같이, RADAR에서는 측위를 위한 장비로 세 개의 기지 스테이션을 두고, 사용자의 위치를 기지 스테이션에서 판정한다.

이상에서 살펴본 옥내 측위 방법들은 측위를 위한 특

수 장비와 환경을 갖추어야 한다. 특수 장비를 갖추려면 상응하는 비용을 투자해야 할 뿐만 아니라 장비가 설치되어 있지 않은 곳에서는 적용할 수 없다는 단점이 있다. 그래서 측위를 위한 특수 장비가 전혀 필요 없는, 무선 랜을 이용하는 방식이 이미 여러 곳에서 발표된 바 있다. 무선 랜을 이용하는 여러 가지 방법 중, RADAR 시스템이 채용한 바와 같은 지문방식을 채용하는 방식이 가장 효율적인 것으로 알려져 있다. 따라서 본 논문은 측위를 위한 특수 장비를 전혀 사용하지 않고 무선 랜을 위하여 기존에 설치된 AP를 사용하면, RADAR 시스템이 채용한 바와 같은 지문방식을 채용하는 옥내 측위 기술들을 살펴보고, 판단나무(decision tree)를 적용한 새로운 지문방식을 제안한다. 그리고 제안한 방법과 기존의 방법의 성능을 비교 분석한다.

II. 본 론

1. 관련연구

옥내 측위 방법에 대한 다양한 연구 결과가 발표되었다. AP (Access Point)와 이동물체와의 거리를 D, 이동물체에서 측정된 AP의 신호세기를 S라 하고, D와 S 사이의 상관관계를 구하여 사용자의 위치를 판정하는 방법의 실험 결과가 [12]에 발표되었다. 그러나 이 방법은 [10]에 언급된 바와 같이 오차가 매우 크다. 이와 관련하여 [10]의 본문에 다음과 같은 문장이 있다:

고주파 신호가 환경의 영향을 너무 크게 받아서 다양한 방법으로 실험을 하여 보았지만 수학적 모델에 맞는 신호 값을 얻기가 불가능하였다.

무선 LAN(Local Area Network)용 AP의 신호 세기를 특징 값으로 하는 지문방식에서 사용하는 기술에는 확률적 방법, K-NN, 그리고 신경망이 있다. 이들 세 가지 기술들의 성능 평가 결과가 [13]에 발표되었다. 본 논문에서는 판단나무 방법을 고안하고, 이론적 그리고 실험적으로 이들 세 가지 기술들과 비교 분석한다.

[14, 15]는 TOA(Time Of Arrival)와 TDOA(Time Difference Of Arrival)와 같은 시간 기반 기술의 정확도를 제고하는 방안을 논의하고, 시뮬레이션으로 제안한 방법의 효율성을 검증하였다. TOA는 이동물체가 신호를 송신한 시각으로부터 기지스테이션에 도달한 시각의 차이로부터 거리를 계산하는데 이동물체의 시계와 기지스테이션의 시계가 동기화되기 어렵기 때문에 정확한 거리를 구하는데 어려움이 있다. TDOA는 여러

개의 기지 스테이션들이 각각 도달시각을 측정하여 이들 간의 차이를 바탕으로 거리를 계산한다. 거리를 구한 다음에는 삼각측량법으로 이동물체의 위치를 계산한다. 제안하는 방법들은 특수 장치가 필요하며 아직 구현이 불가능함으로 시뮬레이션으로 제안하는 방법을 검증하였다.

[16]은 네트워크 기반 무선 측위 기술들로 TOA와 TDOA와 같은 시간 기반 기술들을 설명하고 간섭이나 감쇄와 같은 이 분야에서 극복해야 할 과제들을 개괄적으로 살펴본다. [17]은 사용자의 위치를 이산적인 정보가 아니라 연속적인 정보로 보고 판정하는 질량의 중심(Center-of-Mass)이라는 방법과 시간-평균(Time-Averaging)이라는 두 가지 방법을 소개한다. 질량의 중심 방법은 모든 후보지점들에 대하여 사용자가 현재 그곳에 위치할 확률을 구하고 확률 값이 가장 큰 N 개의 지점들의 평균을 사용자의 위치로 판정한다. 시간-평균은 질량의 중심 방법으로 구한 최근 W 개의 추정 값의 평균을 현재 위치로 판정한다.

[18]은 후보지점들을 신호가 잡히는 AP들을 고려하여 무리를 지음으로써 실시간 측정 단계에서 비교할 후보지점들의 대상을 축소하는 방법을 소개한다. [19]는 무선 LAN의 AP의 신호세기를 바탕으로 위치를 판정하는 방법의 특성으로 자신의 위치를 노출하지 않는다는 것, AP 중 몇 개가 작동을 하지 않더라도 측위가 가능하다는 것, 등을 들고, 찾아보기 표를 작성할 때 사용한 샘플들의 수와 샘플들의 표준편차를 가중치로 사용하는 K-NN 알고리즘을 사용하여 후보지점들 간의 거리와 오차의 관계를 살펴보는 실험을 수행하였다.

[20]은 계층적 베이저안 그래프 모형을 측위에 최초로 적용하였으며 여러 사용자의 위치를 동시에 찾을 수 있는 방법을 제안한다. [21]은 혼합 측위 방법을 제안한다. 이 방법은 신호세기와 거리와의 상관관계식을 이용하여 사용자의 위치를 판별하는데, 사용 중에 찾아보기 표도 작성한다. 이 방법의 특징은 찾아보기 표를 반영하여 상관관계식을 무단히 개선하는 것이다.

[22, 23]은 찾아보기 표를 작성하는 수고를 더는 방법을 제안한다. [22]는 확률적 방법으로 사용자 영역을 물리적 구획으로 나누는 방법을 제안하고, [23]은 라디오 파장의 전파 모형을 이용하여 찾아보기 표를 작성하는 방법을 제안한다.

[24, 25]는 K-NN 방법을 분석하는 프레임워크를 제안하고 이 프레임워크를 이용하여 K-NN 방법을 사용할 때 신호가 잡히는 AP의 수는 다섯 개가 적당하며,

후보지점간의 거리는 1m가 적당하다는 것을 비롯하여 측위에 영향을 주는 요소들을 분석한다. 또한 이 프레임워크를 이용하여 측위 환경을 최적화하도록 관련 인자의 값을 조정하는 방법도 [25]에 소개된다.

2. 지문방식 알고리즘들

지금까지 발표된 지문방식 알고리즘에는 K-NN 방법, 베이저안(Bayesian) 방법, 신경망 방법, 등이 있다. K-NN 방법은 1절에서 소개한 RADAR 시스템이 채용한 방식이다.

베이저안 방법은 실시간 측정 단계에서 채취한 지문(일련의 AP의 신호 세기들로 구성된 배열)이 주어진 조건에서 각 후보지점이 발생할 조건부 확률이 가장 큰 후보지점을 현재 사용자의 위치로 판단하는 측위 방법이다.

예를 들어, 실시간 측정 단계에서 실측하여 얻은 AP들의 신호 세기들로 구성된 배열이 $X = (x_1, x_2, \dots, x_n)$ 라 하고, 후보지점의 수를 m이라 하자. 다음 식을 만족하는 후보지점 CP_i 를 사용자의 현재지점이라고 베이저안 알고리즘은 결정한다.

$$P(CP_i|X) > P(CP_j|X) \text{ for } 1 \leq j \leq m, j \neq i \quad (1)$$

베이의 정리에 의하여 다음이 참이다:

$$P(CP_i|X) = \frac{P(X|CP_i)P(CP_i)}{P(X)}$$

모든 후보지점에 대하여 $P(X)$ 는 상수임으로 $P(X|CP_i)P(CP_i)$ 를 최대로 하여주는 CP_i 가 곧 (식 1)을 만족하는 후보지점이 된다. 따라서 베이저안 알고리즘은 $P(X|CP_i)P(CP_i)$ 를 극대화하는 CP_i 를 찾아준다^[26].

신경망은 여러 층으로 구성되고, 각 층은 여러 개의 노드(신경이라고도 함)로 구성된다. 어떤 층의 신경은 이웃 층의 신경들과 간선으로 연결되어 있고, 각 간선에는 무게(weight)가 연합되어 있다. 또한 각 신경에는 한계값(threshold)도 연합되어 있다. 각 신경은 이전 층의 신경이 출력하는 값과 간선의 무게를 곱한 값들의 총합과 한계값으로 출력값을 산출한다. 어떤 신경의 출력값은 다음 층의 신경의 입력값이 된다. 마지막 층의 신경은 후보지점을 나타내며, 이들의 출력값으로 어느 후보지점이 사용자의 현재 지점인지를 판단한다. AP의 수가 n개이고 후보지점의 수가 m일 때, 측위 신경망의 첫 번째 층을 구성하는 신경의 수는 n이고 마지막 층을

구성하는 신경의 수는 m 이다. 신경망 관련 논문은 이미 수도 없이 발표된 바 있다^[27, 28].

3. 판단 나무 알고리즘

판단나무 알고리즘은 준비단계에서 훈련 데이터 (training data)로 판단나무를 건축한다. 훈련데이터는 준비단계 동안 각 후보지점에서 실측한 신호세기 배열

표 2. 훈련데이터의 예 (C.P는 후보지점, CPi는 i번 후보지점의 좌표, APi는 i번 AP의 MAC 주소, I는 급간격임)

Table 2. Example of training data set (C.P stands for candidate position, CPi is the coordinates of i-th candidate position, APi is the i-th AP's MAC address, and I stands for interval.)

C.P \ AP	AP1	AP2	AP3	AP4	AP5
CP1	I2	I1	I2	I5	I5
	I1	I1	I2	I1	I1
	...				
CP2	I3	I2	I3	I1	I2
	...				
...	...				

```

algorithm Construct_DT(int[][] Table,
                      ListType MacList, int Index)
(1)  If (Number of rows in Table != 1)
    A. If all the CPs of Table are the same, CPi, then
       Tree[Index] = CPi and return
    B. else if (Number of rows in Table == 0) then
       Tree[Index] = Empty and return.
    C. else if (Number of columns in Table == 1) then
       Tree[Index] = for each CPi in Table, probability
       of CPi and return
(2)  else Tree[Index] = CP and return
(3)  end if
// Number of rows != 1 and !A and !B and !C
(4)  Compute I
(5)  Compute Entropies for each AP
(6)  Tree[Index] = MacAddress of the AP with
       maximum Gain and array P. P[i] is the
       probability of CPi in Table
(7)  Construct subMacList
(8)  loop(i=1; i<=number of Intervals; i++)
    A. generate subTable
    B. subIndex = Index*number of Interval + i
    C. Construct_DT (subTable, subMacList, subIndex)
(9)  end loop
end Construct_DT
    
```

그림 2. 판단나무를 구축하는 알고리즘
Fig. 2. Algorithm to construct a decision tree.

들로 구성되며, 각 후보지점에서 실측하는 회수가 크면 측위가 더욱 정확해지는 경향이 있다. 훈련데이터의 예가 표 2에 보인다.

표 2는 표 1과 대동소이하다. 표 1에서는 신호세기의 평균값을 사용하여 각 후보지점당 하나의 배열이 연합되어 있는데 반하여 표 2에서는 평균값을 쓰지 않고 측정치를 그대로 사용함으로 하나의 후보지점 당 여러 개 (예를 들어 100 개)의 배열이 연합된다. 그리고, 표 1에서는 측정치의 단위를 그대로 사용하는데 표 2에서는 급간격으로 표현한다. 급간격을 결정하는 예로 다음을 들 수 있다.

$$\begin{aligned}
 I1 &= \{x | x > -30\}, \\
 I2 &= \{x | -40 < x \leq -30\}, \\
 I3 &= \{x | -50 < x \leq -40\}, \dots
 \end{aligned}$$

주어진 훈련데이터로 그림 2에 보이는 알고리즘 Construct_DT가 판단나무를 구축한다. 알고리즘의 단계 (4)는 다음과 같은 식으로 표 2와 같은 훈련데이터를 후보지점별로 나누는데 필요한 예상정보량인 I를 구한다.

$$I(s_1, s_2, \dots, s_m) = - \sum_{i=1}^m p_i \log_2(p_i),$$

단, m 은 후보지점의 수, S 는 훈련데이터의 수 (표 2의 행 수), s_j 는 후보지점 CP_i 의 훈련데이터의 수, 그리고 $p_i = \frac{s_i}{S}$ 이다.

알고리즘의 단계 (5)는 엔트로피, 즉 훈련데이터를 k 번째 AP의 신호세기를 기준으로 구분한 결과의 예상정보를 계산한다. AP_k 가 다음과 같이 v (급간격의 수) 개의 값을 갖는다고 하자: $\{a_1, a_2, \dots, a_v\}$. 그러면, AP_k 로 S 를 구분할 때, 다음과 같은 v 개의 부분집합을 얻게 된다: $\{a_1, a_2, \dots, a_v\}$. 단, S_j 는 S 를 구성하는 행 중, AP_k 의 값이 a_j 인 행들의 집합이다. 부분집합 S_j 에서 후보지점 CP_i 의 행의 수를 S_{ij} 라 하자. 그러면, 엔트로피 $E(AP_k)$ 는 다음 식으로 구한다.

$$E(AP_k) = \sum_{j=1}^v \frac{s_{1j} + \dots + s_{mj}}{S} I(s_{1j}, \dots, s_{mj})$$

알고리즘의 단계 (6)은 다음 식으로 정보이득 $G(AP_k)$ 를 구한다.

$$Gain(AP_k) = I(s_1, s_2, \dots, s_m) - E(AP_k)$$

알고리즘 Construct_DT는 정보이득이 가장 큰 AP_k 를 선택하여 훈련데이터를 분류하는 작업을 더 이상 분류가 되지 않을 때까지 반복한다.

4. 시간복잡도분석

준비단계는 실행시간이 결정적인 작업이 아니다. 그러나 실시간 측위 단계는 움직이고 있는 사용자의 위치를 응용프로그램에게 제공하는 과업을 수행함으로 실행시간이 길면 사용자에게 적당한 서비스를 제공할 수 없다. 위에서 살펴본 지문방식 알고리즘들의 실행 시간에 영향을 주는 인자들을 살펴보면 다음과 같다: 급간격의 수 I, AP의 수 N, 그리고 후보지점의 수 M. K-NN 방법은 실시간 측위 단계에서 얻은 실측 벡터 X와 찾아보기 표의 각 행을 비교하기 때문에 실시간 측위 단계의 실행 시간은 $O(N * M)$ 이다.

베이지안 방법의 준비 단계에서 작성한 훈련데이터 표는 표 3과 같은 형태다. 각 후보지점마다 급간격의 수인 I 개의 행을 갖는다. 베이지안 방법의 실측 단계에서 측정된 실측 벡터가 X 일 때, $X(AP_k)$ 는 후보지점 AP_k 에서 잡힌 신호의 세기를 나타낸다. 베이지안 방법은 모든 $X(AP_k)$ 에 대하여 AP_k 의 값이 동일한 훈련데이터의 수를 세기 때문에 베이지안 방법의 시간 복잡도는 $O(I * N * M)$ 이다.

신경망의 실시간 측정 단계의 시간복잡도는 준비 단계에서 구축된 신경망의 층의 수, 노드 수, 그리고 연결 형태에 따라 결정된다. 그런데 측위 신경망의 출력층에는 최소한 M 개의 노드가 있고 각 출력층 노드의 입력 노드의 수는 N이나 I보다 크다. 그래서 출력층의 각 노드는 자신의 출력값을 결정하기 위하여 최소한 N이나 I 개의 연산을 수행해야 한다. 출력층의 노드 수는 M이므로 출력층에서 수행되는 연산에 필요한 시간은 $O(M * I)$ 이거나 $O(M * N)$ 이다. 따라서 신경망 방법의 실시간 측정 단계의 시간복잡도는 $O(M * I)$ 나 $O(M * N)$ 보다 크다.

판단나무의 실시간 단계의 수행시간은 $O(I * N)$ 이다. 그 이유는 다음과 같다. 그림 3의 판단나무의 예에 보이는 바와 같이, 판단나무의 깊이는 N, 즉 AP의 수이고, 각 노드의 가지 수는 급간격의 수, 즉 I이며, 판단나무의 잎사귀노드는 후보지점, 즉 CP_i 이다. 따라서 실시간 측정 단계는 뿌리노드에서 최대 I번 비교를 하여 자

표 3. 베이지안 방법의 준비단계에서 작성한 훈련데이터의 예

Table 3. An example training data set constructed by the Off-line phase of Bayesian Method

	AP1	AP2	AP3	AP4	AP5
CP1	I1: 10	I1: 5	...		
CP1	I2: 9	I2: 7	...		
CP1	I3: 5	I3: 15	...		
CP1	I4: 5	I4: 3	...		
CP1	I5: 1	I5: 0	...		
CP2	I1: 5		
...

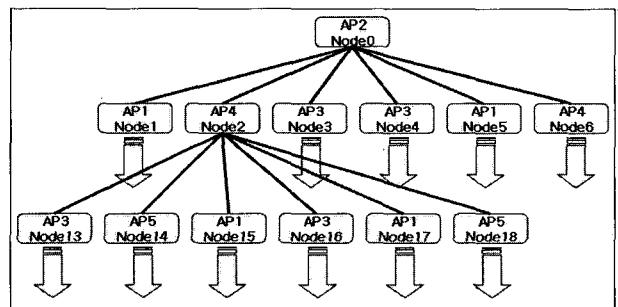


그림 3. 판단나무의 예

Fig. 3. An example decision tree

표 4. 시간복잡도 요약 ('>>')은 훨씬 더 큼을 의미함

Table 4. Summary of time complexities ('>>' means much greater than)

Methods	K-NN	Bayesian	Neural Network	Decision tree
Time Complexity	$O(N * M)$	$O(I * N * M)$	$\gg O(N * M)$	$O(I * N)$

식노드를 결정하고, 자식노드를 뿌리노드로 삼아 이러한 과정을 반복한다. 이 반복은 노드가 후보지점이 될 때까지 즉, 최대 N번 반복한다. 따라서, 실시간 단계의 시간복잡도는 $O(I * N)$ 이다.

시간복잡도 분석 결과가 표 4에 요약되어 있다. 후보지점의 수 M은 일반적으로 N과 I보다 훨씬 크다. 예를 들어, 이 논문의 실험 장소의 경우, I=6, N=5, 그리고 M=96이다. 따라서 판단나무 방법이 다른 방법보다 훨씬 더 효율적이다.

III. 실험

실험적 분석을 위하여 K-NN, 베이지안, 그리고 판단나무 방법들을 삼성 센스 R50 노트북 컴퓨터에 C#을 이용하여 이 알고리즘들을 구현하였다. 노트북 컴퓨터에 장착된 무선랜 카드는 인텔 무선 2200RG이다. 신경

망은 정확성 면에서 K-NN보다 떨어진다고 [13]에 알려져 있어 실험적 분석에서 제외하였다.

1. 실험 장소

실험 장소는 그림 4에 보이는 공학관 4층 마이크로컴퓨터 실험실이다. 각 후보지점 간의 거리는 1m로 하였고, 실험은 오후 8시 이후에 수행하였다. 4층에는 모두 8 개의 AP가 있으나 (그림에서 별표로 표시된 곳에 있음), 다섯 개의 AP 신호만 실험에 사용하였다.

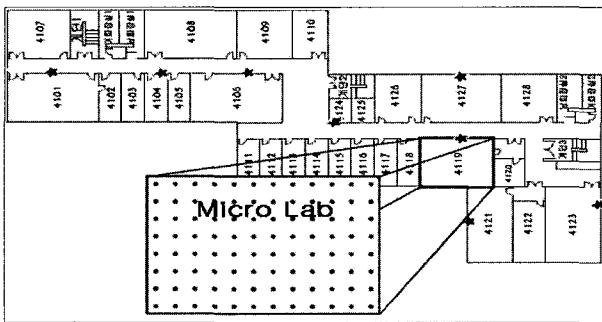


그림 4. 실험 장소인 마이크로 실험실
Fig. 4. Our test bed, Micro Lab.

2. 실험 결과

N=5, I=6, M=96인 경우에 각 후보지점에서 준비단계에 채취한 샘플의 수를 변화하면서 1-NN, 베이지안, 판단나무 방법을 실행하여 정확성을 비교하여 그림 5와 같은 결과를 얻었다. 샘플의 수가 10 개일 때, 1-NN 방법이 다른 방법들보다 훨씬 더 정확하였다. 그러나 그 차이는 샘플의 수가 증가할수록 감소하여 샘플의 수가 50일 때 정확도 면에서 세 가지 방법이 거의 모두 비슷하였다. 실측단계에서 사용하는 실측치 배열 X도 여러 측정치의 평균을 사용할 수 있으며 평균을 구하기 위한 측정회수가 정확도에 영향을 준다. 본 실험에서는 10회 평균을 사용하였다.

IV. 결 론

옥내 측위 방법에 대한 많은 연구 결과가 발표되었다. 그 중에서 무선네트워크를 사용하는 방법은 측위를 위한 특수 장비를 필요로 하지 않기 때문에 최근에 각광을 받고 있다. 무선네트워크를 이용하는 방법 중에는 지문방식이 가장 정확하다고 알려져 있다. 지문방식은 준비단계와 실시간 측위 단계로 구성되는데, 실시간 측위 단계의 실행시간은 적절한 서비스를 제공하기 위한

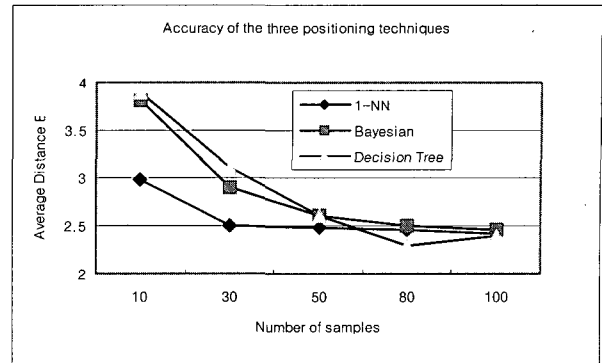


그림 5. 샘플의 수와 세 가지 방법의 정확도의 관계
Fig. 5. Number of samples and the accuracies of the three methods.

매우 중요한 요소임에도 불구하고 실시간 측위 모듈의 효율성을 제고하는 연구가 아직은 없었다.

본 논문은 판단나무 방법이라는 새로운 지문방식 기술을 제안한다. 제안하는 판단나무 방법은 준비 단계에서 판단나무를 구축한다. 그래서 판단나무 방법의 실시간 측위 단계의 시간복잡도는 사용하는 AP의 수가 N이고, 신호세기의 급간격의 수가 I일 때, $O(I*N)$ 이다. 현재까지 소개된 K-NN, 베이지안, 신경망 방법들의 실시간측위 단계의 시간복잡도는, M이 후보지점들의 수일 때, 모두 $O(M*I)$ 나 $O(M*N)$ 보다 훨씬 크다. 일반적으로 M이 I와 N보다 훨씬 크기 때문에 제안하는 판단나무 방법이 기존의 방법보다 효율적임을 알 수 있다.

제안하는 방법과 기존의 방법들과의 정확도를 비교하는 실험을 실시하였다. 실험 결과, 제안하는 방법이 기존의 방법보다 정확도 면에서 결코 뒤지지 않음을 보였다. 결국 본 논문은 기존의 방법보다 정확도가 떨어지지 않으면서 수행 시간이 더 효율적인 판단나무 방법이라는 지문 방식을 위한 새로운 알고리즘을 제안하였다.

향후 연구는 제안한 방법을 이용하여 실용적인 위치 기반서비스 시스템을 구축하는 것이다.

참 고 문 헌

[1] 박종현, 김문구, 백종현, "위치기반서비스(LBS)의 산업구조 분석 및 시장개발전략 방향," 한국통신학회지 Vol. 20, No.4, 2003년 4월.
[2] 주재순, 염경환, 이상정, "GPS/ GLONASS 통합 수신용 RF 전단부의 설계 및 제작," 정보과학회논문지, 제12권 4호, 2001, pp. 1226-2285.

- [3] Bullock, J.B., King, T.M., Kennedy, H.L., Berry, E.D., Zanfino, G., "Test results and analysis of a low cost core GPS receiver for time transfer applications," Proceedings of the 1997 IEEE International Frequency Control Symposium, 1997, 28-30 May 1997 pp.314-322.
- [4] Jefferson, D.C., Lichten, S.M., Young, L.E., "A test of precision GPS clock synchronization," Proceedings of the 1996 IEEE International Frequency Control Symposium, 1996. 50th. 5-7 June 1996 pp.1206-1210
- [5] Nielson, J., Keefer, J., McCullough, B., "SAASM: Rockwell Collins' next generation GPS receiver design," IEEE 2000 Position Location and Navigation Symposium 13-16 March 2000 pp.98-105
- [6] Peczalski, A., Kriz, J., Carlson, S.G., Sampson, S.J., "Military/civilian mixed-mode Global Positioning System (GPS) receiver (MMGR)," Proceedings of the 2004 IEEE Aerospace Conference, 2004, Vol. 4, 6-13 Mar 2004 pp.2697-2703
- [7] Want, R., Hopper, A., Falcao, V., and Gibbons, J., "The Active Badge Location System," ACM Transactions on Information Systems 10, 1 (January 1992), pp. 91-102
- [8] Harter, A. and Hopper, A., "A New Location Technique for the Active Office," IEEE Personal Communications 4, 5 (Oct. 1997), pp. 43-47.
- [9] Harter, A., Hopper, A., Steggles, P., Ward, A., and Webster, P., "The Anatomy of a Context-Aware Application," Proc. 5th ACM MOBICOM Conf. (Seattle, WA, Aug. 1999).
- [10] Priyanthat, N., Chakraborty, A. and Balakrishnan, H., "The Cricket Location-Support System," Proc. of 6th ACM International Conference on Mobile Computing and Networking, Boston, MA, Aug. 2000.
- [11] Bahl, P. and Padmanabhan, V., "RADAR: An in-building RF-based user location and tracking system", INFOCOM 2000, Mar. 2000, pp. 775-784.
- [12] Lassabe, F., Canalda, P., Chatonnay, P., Spies, F., "A Friis-based calibrated model for WiFi terminals positioning," Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. 13-16 June 2005, pp. 382-387
- [13] Lin, T.N. and Lin, P.C., "Performance Comparison of Indoor Positioning Techniques based on Location Fingerprinting in Wireless Networks," Proceedings of the 2005 International Conference on Wireless Networks, Communications and Mobile Computing, Volume 2, 13-16 June 2005, pp. 1569-1574.
- [14] Wann, C., Lin, M., "Data fusion methods for accuracy improvement in wireless location systems," IEEE Wireless Communications and Networking Conference, 2004. (WCNC 2004) Vol. 1, 21-25 March 2004, pp. 471 - 476
- [15] Yousef, N.R., Sayed, A.H., Jalloul, L.M.A., "Robust wireless location over fading channels," IEEE Transactions on Vehicular Technology, Vol. 52, Issue 1, Jan. 2003 pp. 117-126
- [16] Sayed, A.H., Tarighat, A., Khajehnouri, N., "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," IEEE Signal Processing Magazine, Vol. 22, Issue 4, July 2005 pp. 24-40
- [17] Youssef, M., Agrawala, A., "Continuous Space Estimation for WLAN Location Determination Systems, " Proceedings of 13th International Conference on Computer Communications and Networks, 2004 (ICCCN 2004) pp. 161-166
- [18] Youssef, M., Agrawala, A., Shankar, A.U., "WLAN Location Determination via Clustering and Probability Distributions," Proceedings of IEEE International Conference on Pervasive Computing and Communications (PerCom), 2003, 23-26 March 2003, pp. 143-150.
- [19] Prasithsangaree, P., Krishnamurthy, P., Chrysanthis, P., "On indoor position location with wireless LANs," The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002. Vol. 2, 15-18 Sept. 2002 pp.720-724
- [20] Madigan, D., Einahrawy, E., Martin, R.P., Ju, W., Krishnan, P., Krishnakumar, A.S., "Bayesian indoor positioning systems," Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005) vol. 2, 13-17 March 2005 pp. 1217-1227
- [21] Kwon, J., Dunder, B., Varaiya, P., "Hybrid algorithm for indoor positioning using wireless LAN," IEEE 60th Vehicular Technology Conference, 2004. (VTC2004) Fall. 2004 Vol. 7, 26-29 Sept. 2004 pp. 4625-4629
- [22] Roos, T., Myllymaki, P., Tirri, H., Miskangas, P. and Sievanen, J., "A Probabilistic Approach to WLAN User Location Estimation," International Journal of Wireless Information Networks, Vol. 9, No. 3, July 2002.

- [23] Hatami, A., Pahlavan, K., "Comparative statistical analysis of indoor positioning using empirical data and indoor radio channel models," IEEE 3rd Consumer Communications and Networking Conference, 2006 (CCNC 2006) Vol. 2, 8-10 Jan. 2006, pp. 1018-1022
- [24] Kaemarungsi, K., Krishnamurthy, P., "Modeling of indoor positioning systems based on location fingerprinting," Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004), Vol. 2, 7-11 March 2004 pp. 1012-1022
- [25] Kaemarungsi, K., "Efficient Design of Indoor Positioning Systems Based on Location Fingerprinting," 2005 International Conference on Wireless Networks, Communications and Mobile Computing, Vol. 1, 13-16 June 2005 pp.181-186
- [26] Ito, S., Kawaguchi, N., "Bayesian based location estimation system using wireless LAN," Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications Workshops, 2005. (PerCom 2005 Workshops) 8-12 March 2005, pp. 273 - 278.
- [27] Funahashi, K., "On the Approximate Realization of Continuous Mappings by Neural Networks," Neural Networks, 2(3) 1989, pp.183-192
- [28] Hornik, K., Stinchcombe, M, White, H., "Multilayer Feedforward Networks Are Universal Approximators," Neural Networks, 2(5), 1989, pp. 359-336

 저 자 소 개



임 재 결(평생회원)
 1981년 동국대학교 전자계산
 학과 학사
 1987년 University of Illinois
 석사
 1990년 University of Illinois
 박사

1992년~현재 동국대학교(경주) 컴퓨터멀티
 미디어학과 교수

<주관심분야 : 컴퓨터알고리즘, 컴퓨터 시스템 분
 석, 위치기반서비스>