

논문 2011-48SP-2-15

효율적인 실내 측위를 위한 최적화된 KNN/IFCM 알고리즘

(Optimized KNN/IFCM Algorithm for Efficient Indoor Location)

이 장 재*, 송 익 호**, 김 중 화***, 이 성 로***

(Jang Jae Lee, Lick Ho Song, Jong Hwa Kim, and Seong Ro Lee)

요 약

WLAN 환경하에서 알고리즘 기반의 패턴 매칭을 위해 training 단계에서는 여러 개의 AP에서 신호 잡음비의 특성값을 데이터베이스에 만들어 활용하고 estimation 단계에서는 단말기(MU)의 2차원 좌표값을 단말기로부터 새롭게 얻은 SNR과 데이터베이스에 저장된 fingerprint를 비교함으로써 추정한다. Fingerprinting 방식에서 KNN은 WLAN 기반 실내 측위에 가장 많이 적용되고 있지만 KNN의 성능은 k 개의 이웃 수와 RP의 수에 따라 민감하다. 논문에서는 KNN 성능을 향상시키기 위해 PFCM 군집화를 적용한 KNN과 PFCM을 혼합한 알고리즘을 제안하였다. 제안한 알고리즘은 신호잡음비 데이터를 KNN 방법에 적용하여 k 개의 RP를 선택한 후 선택된 RP의 신호잡음비를 PFCM에 적용하여 k 개의 RP를 군집하여 분류한다. 실험 결과에서는 위치 오차가 2m 이내에서 KNN/IFCM 알고리즘이 KNN, KNN/FCM, KNN/PFCM 알고리즘보다 성능이 우수하다.

Abstract

For any pattern matching based algorithm in WLAN environment, the characteristics of signal to noise ratio(SNR) to multiple access points(APs) are utilized to establish database in the training phase, and in the estimation phase, the actual two dimensional coordinates of mobile unit(MU) are estimated based on the comparison between the new recorded SNR and fingerprints stored in database. As fingerprinting method, k -nearest neighbor(KNN) has been widely applied for indoor location in wireless location area networks(WLAN), but its performance is sensitive to number of neighbors k and positions of reference points(RPs). So intuitive fuzzy c-means(IFCM) clustering algorithm is applied to improve KNN, which is the KNN/IFCM hybrid algorithm presented in this paper. In the proposed algorithm, through KNN, k RPs are firstly chosen as the data samples of IFCM based on signal to noise ratio(SNR). Then, the k RPs are classified into different clusters through IFCM based on SNR. Experimental results indicate that the proposed KNN/IFCM hybrid algorithm generally outperforms KNN, KNN/FCM, KNN/PFCM algorithm when the locations error is less than 2m.

Keywords : WLAN, 실내측위, Fingerprinting, k -Nearest Neighbor, Intuitive Fuzzy C-Means

* 정회원, 조선대학교 컴퓨터통계학과

(Dept. of Computer Science and Statistics, Chosun University)

** 평생회원, 한국과학기술원 전기 및 전자공학부

(Dept. of Electrical Engineering, KAIST)

*** 평생회원, 목포대학교 컴퓨터공학과

(Dept. of Computer Engineering, Mokpo National University)

*** 정회원, 목포대학교 정보전자공학과

(Dept. of Information and Electronics Engineering, Mokpo National University)

※ 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 대학중점연구소지원사업으로 수행된 연구임(2009-0093828)

※ 이 논문은 2008학년도 목포대학교 기성회 연구비 지원에 의하여 연구되었음.

접수일자: 2010년10월11일, 수정완료일: 2010년11월24일

I. 서론

최근 새로운 실내 위치 측위 방법으로 무선랜(WLAN : Wireless Local Area Network)을 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^[1].

무선 랜을 이용하는 여러 가지 방법 중 RADAR 시스템이 채용한 바와 같이 fingerprinting 방식을 채용하는 방식이 가장 효율적인 것으로 알려져 있다^[2]. 무선랜을 이용한 위치 측위 결정을 위한 fingerprinting 방식 알고리즘은 KNN(k -nearest neighbor) 방법, 베이저안(bayesian) 방법, 신경망(neural network), 판단 나무(decision tree) 방법^[3] 등이 있다. 이들 알고리즘은 학습 데이터의 각 개체가 특정 클래스에 속할 확률을 예측해주는 통계적 알고리즘이기 때문에 유사데이터 중 분류 모형 속의 데이터를 예측하는데 주로 사용한다.

Fuzzy C-Means(FCM) 군집화 방법은 하나의 클러스터에 속해져 있는 각각의 데이터들을 클러스터에 대한 데이터의 소속정도에 따라 분류하는 데이터 분류 알고리즘으로 Bezdek에 의해 제안되었다^[4]. FCM은 여러 분야에서 성공적으로 적용되었지만 FCM은 하나의 포인트가 모든 클러스터에 속할 정도의 합이 1이 되도록 함으로써 때때로 직관적인 분할과는 다른 결과를 보여준다. 또한 이러한 제약 사항으로 인해 FCM은 잡음에 민감한 특성을 갖는다. 이러한 문제를 해결하기 위해 FCM의 제약을 제거한 Possibilistic C-Mean(PCM)을 제안하였다. PCM은 기존의 확률적인 방법이 아닌 가능성 이론을 사용하여 잡음이나 특이점(outlier)들이 어느 클러스터에도 소속되지 않도록 함으로써 잡음 민감성을 제거하였다. 하지만 PCM은 각각의 클러스터들이 서로 영향을 주지 않으므로 초기화에 민감하고 때때로 중복된 클러스터를 찾아내는 문제점이 있다^[5]. 이러한 FCM의 잡음 민감성과 PCM의 중첩 클러스터 문제를 극복하기 위해 확률적 측도 또는 소속도와 가능성 측도 또는 전형성을 함께 사용하는 방안이 연구되었으며^[6] 그 중 하나가 PFCM이다. PFCM은 FCM과 PCM의 장점을 결합하여 잡음에 강하고 클러스터들이 일부 중첩된

경우에도 중심을 정확하게 찾아낼 수 있다. 하지만 PFCM은 파라미터 설정에 따라 민감하고 잡음에 대한 성능의 향상에 따른 계산량이 상대적으로 증가되는 단점이 있다. 이러한 FCM 알고리즘의 잡음 문제를 해결하고 PCM 알고리즘의 초기값 문제, PFCM 알고리즘의 파라미터 설정 문제들에 대해 효과적으로 해결할 수 있는 Intuitive Fuzzy C-Means(IFCM) 알고리즘을 제안하였다^[11]. IFCM 알고리즘은 FCM 알고리즘의 소속도를 대체하는 직관적인 소속도(Intuitive Membership)를 고안하여 효과적으로 군집들을 학습할 수 있다. 또한 잡음이나 다른 군집에 속한 데이터들에 대한 상대적으로 작은 소속도를 부여하기 때문에 각각의 군집을 분류하기 위해 효과적이다.

fingerprinting 방식에 속하는 기술은 다양하게 많이 있는데 그 중에서 가장 기본적으로 사용되는 기술은 KNN 방법이다. KNN은 WLAN 기반 실내 측위에 간단하게 적용가능하고 성능적인 측면에서 가장 많이 적용되고 있지만 KNN은 k 개의 이웃 수와 RP의 수에 따라 민감하고 광범위하고 복잡한 실내환경에서 거리오차가 많이 발생하여 실제 환경에서 최적의 위치를 찾기가 어렵다. 따라서 이러한 문제를 보완하기 위해서 fingerprinting 방식에서 가장 많이 적용되는 KNN 방법으로 분류되어진 k 개 그룹 데이터는 Fuzzy 클러스터링 알고리즘 중에서 가장 성능이 우수한 IFCM 알고리즘을 사용하여 최적의 위치를 식별해내는 접근 방식인 KNN/IFCM 혼합 알고리즘을 적용하면 거리오차를 최소화 할 수 있어 최적의 위치를 찾을 수 있다.

II. 기존의 fuzzy clustering

1. Fuzzy C-Means Algorithm

FCM 군집화 방법은 하나의 클러스터에 속해져 있는 각각의 데이터들을 클러스터에 대한 데이터의 소속정도에 따라 분류하는 데이터 분류 알고리즘으로 Bezdek에 의해 제안되었다. FCM 군집화는 n 의 벡터의 집합을 c 개의 퍼지 그룹으로 분할하고 소속정도 측정 비용함수가 최소가 되는 것과 같은 각각의 그룹 안에서 클러스터의 중심을 찾는다.

FCM 군집화 방법은 0과 1사이의 소속정도에 의해서 나타난 소속감(membership) 정도를 가지고 주어진 데이터 점이 몇 개의 그룹에 속할 수 있다는 퍼지분할을 사용한다. 즉, 퍼지 분할을 적용하기 위해서 소속 함수

U 는 0과 1 사이 값을 가지는 요소들을 가지며 데이터 집합에 대한 소속감 정도의 합은 항상 1이다.

FCM에서 각 SNR 데이터 포인트 x_k 는 각 퍼지 부분집합 즉, 각 클러스터에 소속되는 정도를 가지며, 목적 함수는 식 (1)과 같이 주어진다.

$$J(u_{ik}, v_i) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2 \quad (1)$$

이때 $1 < m < \infty$ 은 fuzziness 정도를 나타내는 상수로 일반적으로 2로 설정된다. d_{ik}^2 는 k 번째 SNR 데이터 포인트와 i 번째 클러스터 중심 사이의 거리 제곱을 나타내며, 라그랑주(lagrange)의 방법을 사용하여 u_{ik} 와 v_i 는 각각 식(2), (3)과 같이 구해지고^[4], 일반적으로 AO(Alternating Optimization)방법을 통해 국부 최대값을 구한다.

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}} \quad (2)$$

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad (3)$$

FCM은 하나의 포인트가 모든 클러스터에 속할 정도의 합이 1이 되도록 함으로써 때때로 직관적인 분할과는 다른 결과를 보여준다. 또한 이러한 제약사항으로 인해 FCM은 잡음에 민감한 특성을 갖는다.

2. Possibilistic C-Means Algorithm

PCM은 FCM의 직관적이지 못한 소속도 값을 개선하기 위해 소속도 값의 합이 1되는 제약 사항을 제거한 것이다. PCM에서는 각 포인트가 클러스터에 속하는 정도를 표현하기 위해 소속도가 아닌 전형성(typicality)을 사용한다. PCM의 목적 함수는 식(4)와 같으며, 식에서 두 번째 항은 모든 t_{ik} 값이 0이 되는 경우 목적 함수가 최소화되는 자명해(trivial solution)를 제거하기 위해 첨가된 항이다.

PCM 군집화에 대한 목적함수는 식 (4)와 같은 형태를 가지고 있다^[5].

$$J(t_{ik}, v_i) = \sum_{i=1}^c \sum_{k=1}^n t_{ik}^m d_{ik}^2 + \sum_{i=1}^c \delta_i \sum_{k=1}^n (1 - t_{ik})^m \quad (4)$$

여기서,

$$d_{ik} = d(x_k - v_i) = \left(\sum_{j=1}^1 (x_{kj} - v_{ij})^2 \right)^{1/2}$$

$$\delta_i = K \frac{\sum_{k=1}^n u_{ik}^m d_{ik}^2}{\sum_{k=1}^n u_{ik}^m}, \quad K > 0 \quad (5)$$

$$v_i = (v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{iL})$$

여기서, t_{ik} 는 0과 1사이의 수적인 값으로 i 번째 클러스터에 속해져 있는 x_k 의 k 번째 데이터의 전형성 정도를 나타낸다. v_i 는 i 번째 클러스터의 중심 벡터이다. δ_i 는 각 클러스터의 부피를 나타내는 값이다. $j(j=1, 2, \dots, L)$ 는 특성 공간상의 변수이다. m 은 전형성 함수의 퍼지성의 정도에 대한 영향을 나타내는 지수의 가중치이다. 이 값은 $m \in [1, \infty)$ 와 같은 범위를 가지고 있으며 분류 과정에서 퍼지성의 양을 제어하는 파라미터이다. 보통 m 을 2로 설정한다.

PCM 역시 FCM과 같이 AO 방법을 통해 국부 최대값을 구하는 것이 일반적이며 t_{ik} 값은 식 (6)과 같이 주어지고 v_i 값은 식 (7)과 같다.

$$t_{ik} = \frac{1}{1 + \left(\frac{d_{ik}^2}{\delta_i} \right)^{1/(m-1)}} \quad (6)$$

$$v_i = \frac{\sum_{k=1}^n t_{ik}^m x_k}{\sum_{k=1}^n t_{ik}^m} \quad (7)$$

식 (6)은 식(2)가 한 데이터 포인트와 모든 클러스터 중심사이의 거리를 고려하는 것과 달리 한 데이터 포인트와 하나의 클러스터 중심 사이의 거리만을 고려한다. 이러한 특징은 클러스터들 사이의 상관관계를 고려하지 않음으로서 중첩된 클러스터 문제를 유발할 수 있다^[9].

3. Possibilistic Fuzzy C-Means Algorithm

FCM과 PCM은 각각의 장점과 단점이 있으므로 서로의 단점을 보완하기 위한 방법으로 소속도와 전형성을 함께 사용하고자 하는 시도가 있어왔고 그 중 하나는 PFCM^[6]이다. PFCM의 목적 함수는 식 (8)과 같다.

$$J(t_{ik}, v_i) = \sum_{i=1}^c \sum_{k=1}^n (au_{ik}^m + bt_{ik}^\eta) d_{ik}^2 + \sum_{i=1}^c \delta_i \sum_{k=1}^n (1-t_{ik})^\eta \quad (8)$$

이때 a , b 는 소속도와 전형성에 대한 가중치 상수이며, η 는 소속도에서 m 과 동일한 역할을 전형성에서 하는 상수이다. 또한 δ_i 는 각 클러스터의 부피를 나타내는 값으로 식 (5)와 같이 계산되고 클러스터의 크기 추정과 특이점(outlier) 판별에 영향을 미친다. PFCM의 u_{ik} 는 식 (2)와 동일하게 계산되며 t_{ik} 와 v_i 는 각각 식(9), (10)로 계산된다.

$$t_{ik} = \frac{1}{1 + \left(\frac{bd_{ik}^2}{\delta_i} \right)^{1/(\eta-1)}} \quad (9)$$

$$v_i = \frac{\sum_{k=1}^n (au_{ik}^m + bt_{ik}^\eta) x_k}{\sum_{k=1}^n (au_{ik}^m + bt_{ik}^\eta)} \quad (10)$$

PFCM은 클러스터 중심의 계산을 위해 소속도와 전형성을 동시에 고려함으로써 FCM의 잡음 민감성을 완화하고 PCM의 중첩된 클러스터 문제를 해결한다.

III. KNN/IFCM Hybrid Algorithm

1. Fingerprinting 방식

Fingerprinting 방식이라고도 불리는 확률적 모델링에 근거한 방식은 노이즈 및 주위 환경 정보를 위치 추적을 위한 정보로 활용하는 방식으로 현재 무선랜 기반의 측위 시스템에서 가장 많이 사용되는 위치 측위 방식이다^[8].

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 추출한 전파 특성값을 저장한다. 이 과정은 측위 대상이 되는 공간의 모든 RP를 측정할 때까지 반복적으로 수행된다.

확률적 모델링에 근거한 방식은 수집된 전파 특성값이 전파 환경까지 포함되어 있기 때문에 기존의 알고리즘에 비해 우수한 위치 정확도를 제공한다. 그러나 확률적 모델링을 이용한 방식은 위치 추정을 수행하기 위

하여 반드시 데이터베이스를 구축하고, 측위 정확도를 향상시키기 위하여 동일한 위치에서 여러 번의 측정 과정을 수행하여야 한다. 그리고 측정해야 하는 곳의 환경이 변할 때마다 새로운 데이터베이스를 다시 구축하여야 하는 문제점을 가진다^[9].

무선랜 기반의 위치 측위 시스템에서는 하드웨어의 추가 없이 알고리즘의 구성만으로 구현이 가능한 확률적 모델링 알고리즘이 보편적으로 사용되고 있으나 정확한 데이터베이스의 구성을 위해 모든 측위 공간에 대한 실제 신호 세기 수집을 필요로 한다.

2. k-Nearest Neighbor

KNN은 기억 기반 추론기법 중에 하나로 패턴 인식에 많이 활용되고 있다. KNN은 estimation 단계에서 측정된 SNR 데이터가 주어졌을 때 유사도 함수에 기반하여 training 단계에서 fingerprinting 방식으로 미리 저장 되어있는 RP(Reference Point)의 SNR 데이터들의 집합에서 측정된 SNR 데이터와 가장 유사한 k 개의 RP의 SNR 데이터들을 선택한다.

RP와 AP의 수를 각각 m 과 n 으로 가정하고 측정된 SNR 데이터와 RP의 SNR 데이터들의 집합과의 유사도는 식 (1)과 같은 거리를 이용하였다.

$$D_{qi} = \left(\sum_{j=1}^n |s_j - S_{ij}|^q \right)^{1/q} \quad i = 1, 2, \dots, m \quad (11)$$

여기서, S_{ij} 는 training 단계에서 저장된 i 번째 RP에서 j 번째 AP의 SNR 데이터, s_j 는 estimation 단계에서 측정된 j 번째 AP의 SNR 데이터이다^[10].

식 (11)을 이용하여 계산된 거리 값들 중에서 가장 가까운 거리 값을 k 개만큼 선정할 수 있고 식 (12)을 이용하여 위치를 추정 할 수 있다.

$$(\hat{x}, \hat{y}) = \frac{1}{K} \sum_{i=1}^K (x_i, y_i) \quad (12)$$

여기서, (x_i, y_i) 는 i 번째 RP의 좌표이고 (\hat{x}, \hat{y}) 는 추정된 위치 좌표이다.

KNN 알고리즘은 식 (11)에서 모수 k 와 q 에 따라 다양한 패턴 결과가 나타난다. 모수 k 는 추정할 좌표에 대한 RP의 수이고 모수 q 는 거리 타입이다. 일반적으로 모수 q 가 1 일 때는 맨하탄 거리이고 모수 q 가 2 일 때는 유클리드 거리이다. 알고리즘의 복잡성과 위치 예

측을 고려했을 때 일반적으로 모수 q 는 1과 2 일 때 거리를 사용한다.

3. Intuitive Fuzzy C-Means Algorithm

IFCM 알고리즘은 FCM 알고리즘의 소속도를 대체하는 직관적인 소속도(Intuitive Membership)를 고안하여 효과적으로 군집들을 학습할 수 있다. 또한 잡음이나 다른 군집에 속한 데이터들에 대한 상대적으로 작은 소속도를 부여하기 때문에 각각의 군집을 분류하기 위해 효과적이다^[7]. IFCM의 목적함수는 식 (13)과 같다.

$$J(I_{ik}, v_i) = \sum_{i=1}^c \sum_{k=1}^n (I_{ik})^m \|x_k - v_i\|^2 \quad (13)$$

기본적으로 FCM 알고리즘의 목적함수와 동일하나 소속도 대신 직관적인 소속도(I_{ik})를 사용한다. I_{ik} 는 기존의 소속도를 이용하여 식 (14)와 같이 계산된다.

$$I_{ik} = \sum_{i=1}^c (u_{ik})^\eta |u_{ik} - u_{lk}| \quad (14)$$

여기서, $1 \leq i \leq c, 1 \leq k \leq n$ 이다.

I_{ik} 는 식 (14)를 통해 잡음(Outlier)에 대한 높은 소속도 값을 부여했던 FCM 알고리즘은 개선하기 위해 고안되었으며 각 군집에 대해 서로 배타적인 성격을 가지게 되어 직관적인 분류 효과를 가져 올 수 있다. 또한 η 를 이용하여 소속도의 가중치 정도를 선택 할 수 있다. 이때 η 는 소속도에 대한 가중치 상수이고 소속도 함수 u_{ik} 는 식 (2)와 같다.

IFCM 알고리즘의 대푯값을 학습하기 위해 식 (15)을 사용하며 식 (15)는 아래와 같이 유도된다.

$$\begin{aligned} J(I_{ik}, v_i) &= \sum_{i=1}^c \sum_{k=1}^n (I_{ik})^m \|x_k - v_i\|^2, \quad 1 \leq i \leq c \\ \frac{\partial J(I_{ik}, v_i)}{\partial v_i} &= \sum_{i=1}^n (I_{ik})^m (-2)(x_k - v_i) = 0 \\ \sum_{i=1}^n (I_{ik})^m x_k - \sum_{i=1}^n (I_{ik})^m v_i &= 0 \\ \therefore v_i &= \frac{\sum_{i=1}^n (I_{ik})^m x_k}{\sum_{i=1}^n (I_{ik})^m}, \quad 1 \leq i \leq c \end{aligned} \quad (15)$$

(표 1)은 IFCM 알고리즘의 pseudo code 이다. IFCM 알고리즘은 PCM이나 PFCM 알고리즘과 같이 초기화

표 1. IFCM의 pseudo-code
Table 1. Pseudo code of IFCM.

```

Algorithm IFCM
Procedure main()
  Read  $c, m, \epsilon$ 
  [c: initialize cluster,
   m: weighting exponent( $m \in 1, \dots, \infty$ )]
  error=0
  While (error> $\epsilon$ )
    While (input file is not empty)
      Read the data  $x$ 
      [Update IFCM Membership]
      
$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}}$$

       $1 \leq i \leq c, 1 \leq k \leq n$ 
      [Calculate IFCM Intuitive Membership]
      
$$I_{ik} = \sum_{i=1}^c (u_{ik})^\eta |u_{ik} - u_{lk}|$$

       $1 \leq i \leq c, 1 \leq k \leq n$ 
      [Update IFCM Center Mean]
      
$$v_i = \frac{\sum_{i=1}^n (I_{ik})^m x_k}{\sum_{i=1}^n (I_{ik})^m}, \quad 1 \leq i \leq c$$

       $e := v(n+1) - v(n)$ 
    End While
    error = e
  End While
End Main()
End
    
```

에 민감하지 않고 대용량 데이터에 이용가능하며 파라미터 설정 값들이 존재 하지 않으므로 효과적으로 데이터를 학습하고 결과 값을 도출 할 수 있다.

IFCM 알고리즘은 FCM 알고리즘의 잡음 문제를 해결하고 PCM 알고리즘의 초기값 문제, PFCM 알고리즘의 파라미터 설정 문제들에 대해 효과적으로 해결할 수 있다.

4. KNN/IFCM 알고리즘

무선 환경 하에서 패턴 매칭 알고리즘을 위해 training 단계에서 여러 개의 AP로부터 수신된 SNR 값을 측정하여 fingerprinting 방식의 데이터베이스를 만들어 활용해서 측위를 위한 특수 장비를 전혀 사용하지 않고 기존에 설치된 AP를 사용하면서 KNN과 IFCM 군집화 기반의 혼합된 알고리즘을 이용한 무선랜 실내 측위를 결정하는 절차는 (그림 1)과 같다.

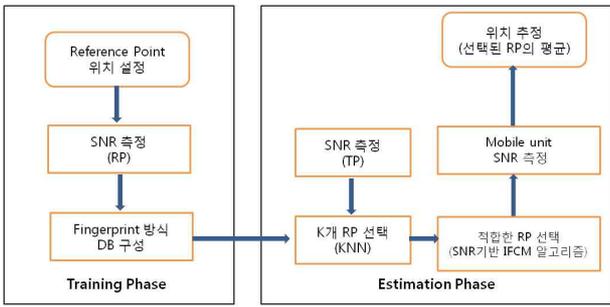


그림 1. KNN/IFCM 혼합 알고리즘 흐름도
Fig. 1. Flow chart of the KNN/IFCM hybrid algorithm.

무선랜 실내 측위 결정을 위해 training 단계와 estimation 단계로 구분한다. training 단계에서는 각 RP에서 수집된 SNR 데이터를 fingerprinting 방식으로 DB화 한다. estimation 단계에서는 KNN/IFCM 혼합 알고리즘을 사용하여 무선랜 실내 위치를 추정하고 결정한다. 첫 번째로 training 단계에서 측정된 SNR 데이터를 KNN 방법을 이용하여 k 개의 RP를 선택한다. 두 번째로 선택된 k 개 RP의 SNR 데이터를 IFCM방법을 이용하여 군집화한다. SNR 데이터에 대한 거리 제곱합이 최소인 클러스터에 속한 RP는 TP(Test Point)에서 측정된 SNR 데이터와 가장 유사한 특징을 갖는다고 가정하면 선택된 RP를 사용하여 실내 위치 추정의 성능을 향상시킨다. 세 번째로 두 번째 단계에서 선택된 군집 중에서 MU에서 측정된 SNR 데이터와 가장 적합한 군집을 선택한 후 군집 내에 있는 RP들의 평균을 산출하여 위치를 추정한다.

III. 실험 및 결과

1. 실험 환경

실험 장소는 (그림 2)에서 보이는 목포대학교 대외협력관 4층 복도이다. 총길이 25m×4m인 복도에 1m 간격으로 63개의 RP 포인트와 40개의 TP 포인트를 (그림 2)와 같이 바둑판 모양으로 지정하였고 RP 포인트는 검정색으로 TP 포인트는 빨간색으로 표시하였다. 2개

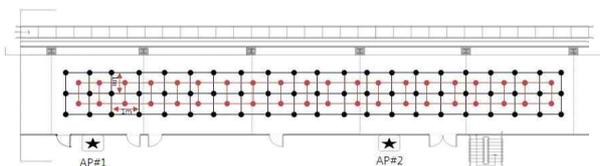


그림 2. 실험 환경
Fig. 2. Experimental environment.

의 AP로부터 1초 간격으로 RP 포인트에서 100회 반복하여 측정하였고 제안한 알고리즘의 성능을 평가하기 위한 TP 포인트에서 60회 반복하여 측정하였다.

2. 실험 장비

실험 장비는 Samsung SENS Q70, NetStumbler 프로그램, AP 2개를 이용하였고 (그림 3)은 무선랜 신호 강도를 측정하는 NetStumbler 프로그램이다. 2개의 AP로부터 각 RP포인트에 수신된 SNR 값을 연속된 그래프와 값으로 측정되었으며 1초 간격으로 SNR 값을 구하여 fingerprinting 방식으로 데이터베이스화 하였다.

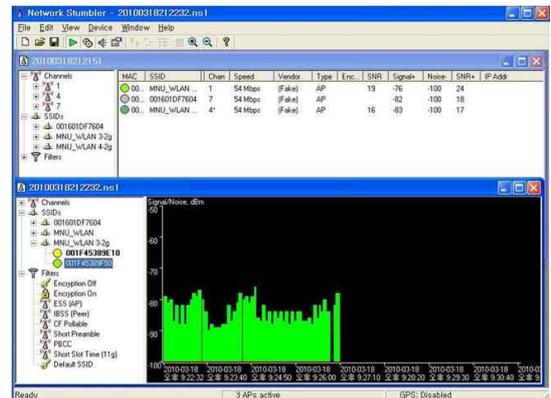


그림 3. NetStumbler
Fig. 3. NetStumbler.

3. 실험 결과

식 (1)에서 모수 k 와 q 에 따라 KNN 알고리즘의 성능 결과가 다양하게 나타난다. 일반적으로 모수 q 가 1 일 때는 맨하탄 거리이고 모수 q 가 2 일 때는 유클리드 거리이다. (그림 4)는 알고리즘의 복잡성과 위치 예측을 고려했을 때 q 는 1과 2를 사용하므로 모수 k 에 따라 평균 거리 오차를 나타낸 결과이다.

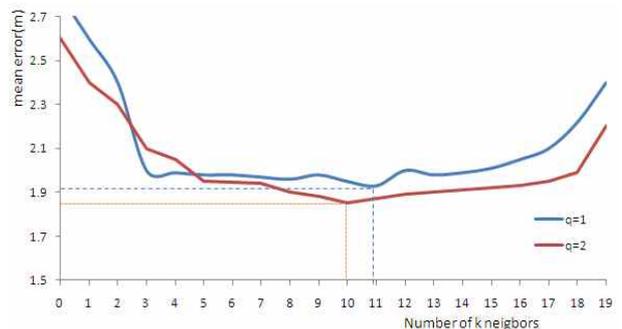


그림 4. k 개의 이웃 수와 평균 거리 오차
Fig. 4. Mean error with different k .

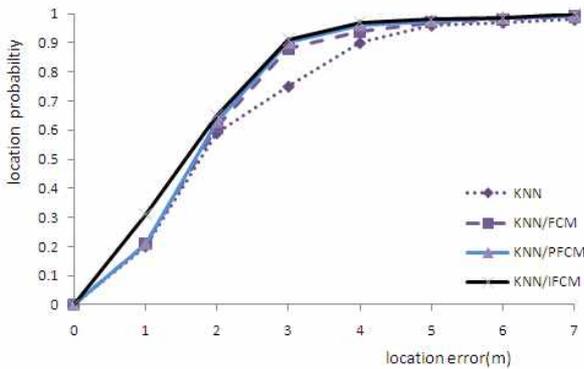


그림 5. 알고리즘의 성능비교($k=10, q=1$)
 Fig. 5. Performance comparison of the algorithms ($k=10, q=1$).

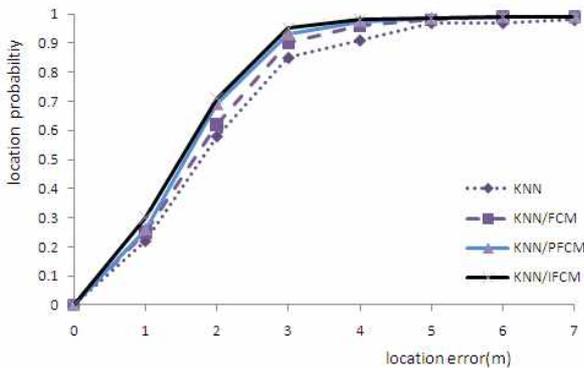


그림 6. 알고리즘의 성능비교($k=10, q=2$)
 Fig. 6. Performance comparison of the algorithms ($k=10, q=2$).

$q=1$ 일 때 모수 k 에 따른 평균 거리 오차에서 k 가 11일 때 평균 거리 오차가 1.93m로 가장 작았고 $q=2$ 일 때 k 에 따른 평균 거리 오차에서 k 가 10일 때 평균 거리 오차가 1.85로 가장 작았고 q 가 1인 맨하탄 거리를 이용하는 KNN 알고리즘보다 q 가 2인 유클리드 거리를 이용하는 KNN 알고리즘이 오차를 줄일 수 있다.

대의 협력관 4층 복도에서 실험한 SNR에 대한 KNN 알고리즘은 $k=10, q=2$ 일 때 최적의 결과가 나타난다. (그림 5)와 (그림 6)은 KNN, KNN/FCM, KNN/PFCM 과 KNN/IFCM 혼합 알고리즘을 이용한 위치 추정 오차와 누적 확률에 관한 결과이다. $k=10, q=1$ 일 때 각 RP에 채취한 SNR을 기반으로 실험한 결과 위치 추정 오차가 2m 이내에서 세 가지의 알고리즘의 누적 확률이 비슷하지만 2m-5m에서 KNN 알고리즘과 KNN/FCM 혼합 알고리즘에 비해 KNN/PFCM 혼합 알고리즘, KNN/IFCM 혼합 알고리즘의 성능이 우수하다고 할 수 있다.

$k=10, q=2$ 일 때는 위치 추정 오차가 2m 이내에서 KNN 알고리즘과 KNN/FCM 혼합 알고리즘에 비해 KNN/PFCM 혼합 알고리즘과 KNN/IFCM 혼합 알고리즘이 더 성능이 우수하고 KNN/FCM, KNN/PFCM과 KNN/IFCM 혼합 알고리즘의 비교에서는 KNN/FCM 혼합 알고리즘에 비해 KNN/PFCM과 KNN/IFCM 혼합 알고리즘의 누적확률이 수치상으로 약간 높으므로 성능이 약간 우수하다고 할 수 있다. 또한 위치 추정 오차가 2m-5m에서도 KNN/IFCM 혼합 알고리즘의 누적확률이 다른 알고리즘보다 성능이 다소 우수하다고 할 수 있다.

KNN 알고리즘은 광범위하고 복잡한 실내 환경에서 계산량이 많기 때문에 실제 환경에서 최적의 위치를 찾기가 어렵다. 따라서 KNN/IFCM 혼합 알고리즘을 적용하면 이러한 문제점을 해결 할 수 있을 것이다. $k=10, q=2$ 일 때 KNN/IFCM 혼합 알고리즘을 적용한 실내 측위 결정이 위치 추정 오차에 상관없이 KNN 알고리즘, KNN/FCM 혼합 알고리즘, KNN/PFCM 혼합 알고리즘보다 더 우수하다.

IV. 결 론

실내 측위 방법 중에서 무선 네트워크를 사용하는 방법은 측위를 위한 특수 장비를 필요로 하지 않기 때문에 최근에 각광을 받고 있다. 무선 네트워크를 이용하는 방법 중에는 fingerprinting 방식이 가장 정확하다고 알려져 있다. 무선 환경 하에서 패턴 매칭 알고리즘을 위해 training 단계에서 여러 개의 AP로부터 수신된 SNR값을 측정하여 fingerprinting 방식의 데이터베이스를 만들어 활용해서 측위를 위한 특수 장비를 전혀 사용하지 않고 무선 랜을 위하여 기존에 설치된 AP를 사용하면서 KNN과 IFCM 군집화 기반의 혼합된 무선 실내 측위 방법을 제안한다. KNN은 k 개의 이웃 수와 RP의 수에 따라 민감하고 광범위하고 복잡한 실내환경에서 거리오차가 많이 발생하여 실제 환경에서 최적의 위치를 찾기가 어렵다. 따라서 이러한 문제를 보완하기 위해서 fingerprinting 방식에서 가장 많이 적용되는 KNN 방법으로 분류되어진 k 개 그룹 데이터는 IFCM 알고리즘을 사용하여 최적의 위치를 식별해내는 접근 방식인 KNN/IFCM 혼합 알고리즘을 적용하면 거리오차를 최소화 할 수 있어 최적의 위치를 찾을 수 있다. $k=10, q=2$ 일 때 KNN/IFCM 혼합 알고리즘을 적용한

실내 측위 결정이 다른 방법들보다 더 우수하다.

향후 연구에서는 Fuzzy 클러스터링 방법들은 기본적으로 사용하는 유클리드 거리 대신에 다양한 거리 측도 방법을 사용한 KNN과 IFCM 혼합된 알고리즘을 이용하여 더 정확한 위치 측위 결과를 얻을 수 있는 연구가 기대된다.

참 고 문 헌

- [1] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, J. Sievanen, "A Probabilistic Approach to WLAN User Location Estimation", *Int. J. Wireless Inform. Network*, Vol. 9, 2002.
- [2] P. Bahl and V.N. Padmanabhan, RADAR: An In-Building RF-based User Location and Tracking System, *Proc. IEEE Computer and Communications Societies*, Vol. 2, 2000.
- [3] 임재철, "옥내 측위를 위한 지문방식 알고리즘들의 성능분석", *전자공학회논문지*, 제43권 CI편, 제6호, 2006년.
- [4] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum, 1987.
- [5] R. Krishnapuram, J. M. Keller, "A Possibilistic Approach to Clustering", *IEEE Transactions on Fuzzy system*, 1993.
- [6] N. R. Pal, K. Pal, J. M. Keller and J. C. Bezdek, "A Possibilistic Fuzzy c-Means Clustering Algorithm", *IEEE Transactions on Fuzzy system*, 2005.
- [7] D. C. Park, "Intuitive Fuzzy C-Means Algorithm", in *Proc. of 2009 IEEE Conf. on Signal Processing and Information Technology*, 2009.
- [8] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughe, F. Potter, J. Tabert, P. Powledge, G. Borriello, B. Schilit, "Place Lab: Device Positioning using Radio Beacons in The Wild", *Pervasive Computing*, Vol. 3468, 2005.
- [9] 김학용, "무선랜 기반 위치정보 서비스", *Telecommunication Review*, 16(4), pp. 580-590 2006.
- [10] A. Hatami, K. Pahlavan, "Comparative Statistical Analysis of Indoor Positioning Using Empirical Data and Indoor Radio Channel Models", *IEEE CCNC 2006 proceedings*. 2006.

저 자 소 개



이 장 재(정회원)
 2000년 조선대학교 전산통계학과 졸업
 2002년 조선대학교 전산통계학과 석사
 2007년 조선대학교 전산통계학과 박사

2009년~2010년 목포대학교 정보산업연구소 전임 연구원
 2011년 현재 조선대학교 컴퓨터통계학과 전임 연구원
 <주관심분야 : 센서네트워크, 실내 위치측위 시스템, 데이터 마이닝, 페이지안 추론>



김 종 화(평생회원)
 1983년 조선대학교 전자공학 학사 졸업.
 1986년 조선대학교 전자공학 석사 졸업.
 1989년 일본 동북대학교 전자공학 박사 졸업.

1991년~ 현재 목포대학교 컴퓨터공학과 교수.
 <주관심분야 : USN, 임베디드시스템, 디지털통신, 해양텔레매틱스>



송 익 호(평생회원)
 1982년~1984년 서울대학교 전자공학과 공학사 (준최우등), 공학석사
 1985년~1987년 펜실베니아대학교 전기공학과 공학석사, 공학박사

1987년~1988년 벨 통신연구소 연구원
 1988년 3월~현재 한국과학기술원 전기및 전자공학과 조교수, 부교수, 교수
 1995년 1월~현재 한국통신학회 논문지 편집위원, 편집 부위원장, 이사
 2010년 1월~현재 한국정보전자통신기술학회 논문지 편집위원, 부회장
 대한전자공학회, 한국음향학회, 한국통신학회 평생회원, IET 석학회원, IEEE 석학회원

<주관심분야 : 통계학적 신호처리와 통신이론, 신호검파와 추정, 이동통신>



이 성 로(정회원)-교신저자
 1987년 고려대학교 전자공학과 졸업
 1990년 한국과학기술원 전기 및 전자공학과 석사
 1996년 한국과학기술원 전기 및 전자공학과 박사

2005년~현재 목포대학교 정보전자공학전공 교수
 <주관심분야 : 디지털통신시스템, 이동 및 위성통신시스템, USN/텔레매틱스응용분야, 임베디드시스템, 생체인식시스템>