

스마트 폰 기반의 가정환경 내 사용자 공간 위치 예측 기법

안 다 예*, 하 란*

Indoor Localization Methodology Based on Smart Phone
in Home Environment

Daye Ahn*, Rhan Ha*

요 약

유비쿼터스 환경에서 실내 공간의 사용자 위치정보는 다양한 응용분야에서 사용자에게 특화된 서비스를 제공하는 데 필요한 필수적인 정보이기 때문에 매우 중요하다. 기존연구들은 규모가 큰 건물에서의 사용자 위치 예측만 고려하고 있고 실험 대상이 되는 공간에서 고정된 AP가 다수 존재한다고 가정한다. 그러나 일반 가정은 면적이 좁은 공간들로 구성되며 고정된 AP가 소수이고 변화가 유동적인 환경이다. 본 논문에서는 기존 연구들이 AP환경이 비교적 안정적인 큰 건물에서의 사용자 위치 예측에 집중한 것과 달리, 일반 가정환경에서 와이파이 핑거프린트 방식을 기반으로 하여 공간을 식별하고 사용자의 위치를 Room-level로 예측하는 사용자 공간 예측 시스템을 제안한다. 실제 가정에서 실험을 한 결과 제안하는 시스템이 모든 가정에서 평균 80%이상의 정확도로 사용자가 위치한 공간을 예측함을 알 수 있었다.

Key Words : Indoor Localization, Wi-Fi Fingerprint, Topic Modeling, GMM, Smartphone

ABSTRACT

In ubiquitous environment, User's location information is very important to serve personalized service to user. Previous works consider only User's locations in the big buildings and assume APs are fixed. Normal home environment, However, is consists of small spaces. And the state of APs is highly fluid. Previous research has focused on indoor localization in the building where has stationary AP environment. However, in this paper, we propose as User's Location Predicting System that finds out a space where a user is located based on Wi-Fi Fingerprint approach in home environments. The results that conducted real home environments are using the system show more than 80% accuracy.

I. 서 론

실내공간의 사용자 위치정보는 에너지 관리를 위한 스마트 홈, 홈 메디컬 등 다양한 분야에서 사용자에게 특화된 서비스를 제공하는데 필요한 사용자 컨텍스트

정보 중 하나이다. 특히 가정환경에서의 에너지 관리를 위해서는 실내 공간, 사용자, 전자제품 사이의 관계를 유추하는 것이 필수적이고 이를 위해서 가장 먼저 선행되어야 할 것은 실내 공간에서의 사용자 위치를 예측하는 것이다. 여기서 실내 공간의 사용자 위치

* This work was supported by the National Research Foundation of Korea grant funded by the Korean government, Ministry of Education, Science and Technology (No.2013-023635).

◆* First Author and Corresponding Author : Hongik University, Dept. of Computer Engineering, Real-time System Lab., dayeahn@mail.hongik.ac.kr, 학생회원

* Hongik University, Dept. of Computer Engineering, Real-time System Lab., rhanha@hongik.ac.kr

논문번호 : KICS2013-12-549, Received December 24, 2013; Revised March 14, 2014; Accepted April 4, 2014

는 사용자가 어떤 독립적인 기능을 갖는 공간에 위치하는지 예측하는 것을 의미한다. 회사건물을 가정해 보면 한 사원이 사무실에 있는지, 회의실에 있는지 아니면 사강실에 있는지 등의 정보를 예측하는 것이라 할 수 있다. 위치정보를 이용하면 독립적인 기능을 갖는 여러 공간과 그 공간에 위치하는 사용자들 사이의 관계를 유추 할 수 있는 등 다양하게 활용될 수 있다. 유비쿼터스를 실현하기 위해 사용자 위치정보를 얻는 것은 매우 중요하기 때문에 최근 실내 공간에서 사용자 위치를 예측하는 연구가 활발하게 진행되고 있다.

실내에서 사용자 위치를 예측하는 기존 연구들은 크게 3가지 유형으로 구분할 수 있다. 첫 번째 유형은 RFID^[1,2], 초음파^[3], 적외선^[4] 등과 같은 기반시설을 추가로 설치해서 얻은 정보를 기반으로 사용자의 위치를 예측하는 기법이다^[5]. 비교적 쉽게 사용자 위치를 예측할 수 있다는 장점을 갖지만 추가적인 기반시설이 필요하기 때문에 비용이 비싸고 시스템의 관리, 유지 및 확장이 어렵다는 특징을 갖는다. 두 번째 유형은 RF 전파 모델을 구축하여 거리에 따른 RF신호의 변화를 하나의 모델로 구축하여 사용자의 위치를 예측하는 기법이다^[6-9]. 다른 유형과 다르게 관심이 있는 모든 지점에서 RF 신호를 수집할 필요가 없다는 장점이 있으나 RF 신호는 주변 환경에 크게 영향을 받기 때문에 기존의 다른 방식들에 비해서 정확도가 떨어지는 단점이 존재한다. 마지막 유형은 RF 핑거프린트를 이용하여 주어진 문제를 해결하는 기법^[8-13]이다. 관심이 있는 위치에서 RF신호를 수집하여 각 공간을 대표하는 값을 추출해 데이터베이스를 구축한 뒤 대표 값으로 실내공간에서 사용자 위치를 예측한다. 특히 RF 핑거프린트를 이용한 유형에서 와이파이 신호를 기반으로 핑거프린트를 수집하여 위치를 예측하는 연구가 활발하게 진행되고 있다. 기존에 존재하는 AP를 이용하기 때문에 첫 번째 유형인 추가적인 기반 시설을 이용하는 방식에 비해 비용이 저렴하지만 각 공간을 대표하는 특징을 추출하기 위해 모든 위치에서 와이파이 신호를 수집해야 하는 특징을 가지고 있다.

위에서 설명한 세 가지 유형에 속하는 연구는 실내 공간에서 사용자의 위치(또는 평면상의 좌표)를 예측하는 것으로, 어떤 공간에서 사용자가 위치하는 지점을 예측하는 것^[14-16]이다. 사용자가 위치하는 지점 예측은 오래 전부터 활발하게 연구되어왔던 분야로 중요한 정보이다. 그러나 최근에 사용자가 위치하는 지점을 예측하는 것과 다르게 사용자가 어떤 독립적인 기능을 갖는 공간에 위치하고 있는지, 즉, Room-level

에서 사용자가 속하는 공간이 어떤 공간인지를 예측하는 연구가 진행되고 있다. RF 핑거프린트를 이용한 유형에서도 특히 와이파이 핑거프린트를 이용하여 사용자 위치를 예측하는 연구 중에는 사용자 위치를 위치하는 지점이 아니라 위치하는 공간과 같이 Room-level로 예측하는 연구들이 있다. 그 중에 한 연구^[17]는 n-gram 알고리즘을 이용하여 와이파이 신호의 잡음을 제거하고 스마트 폰의 가속도 센서를 이용한 사용자의 움직임 정보를 이용하여 높은 정확도로 별도의 학습과정 없이 사용자가 위치하는 공간을 예측한다. 비슷한 접근방법의 또 다른 연구^[18,19]는 n-gram 알고리즘을 이용한 연구^[17]와 비슷한 접근방식을 취하고 있지만 해당 건물의 설계도를 가지고 있다는 가정과 와이파이 핑거프린트를 이용하여 구축한 그래프로 사용자가 위치하는 공간을 예측한다는 점에서 차이가 있다. n-gram 알고리즘을 이용한 연구와 그래프를 이용한 연구 모두 사용자 위치를 Room-level로 예측하는데 있어 높은 정확도를 보인다는 점에서 바람직하다. 그러나 두 연구에서는 모두 와이파이 핑거프린트로 사용자 위치를 예측하기 위해 사용하는 AP들의 정확한 위치를 알고 있고 AP들의 위치가 고정되어있다고 가정한다. 그러나 상대적으로 크기가 작고 각 공간들이 명확하게 구분되지 않는 일반 가정은 대형 건물에 비해 설치된 AP 성능이 떨어지며 수집되는 AP의 개수나 세기와 설치된 위치가 유동적인 환경이다. 따라서 일반 가정환경에서 사용자 위치를 예측하기 위해서는 유동적인 AP환경의 변화에도 적응할 수 있는 다른 접근방법이 필요하다.

본 논문에서는 공간의 넓이가 좁고 공간 간의 경계가 명확하게 구분되지 않는 일반 가정환경에서 와이파이 핑거프린트 기법을 기반으로 사용자가 위치하는 공간을 예측하는 시스템을 제안한다. 일반 가정환경은 AP환경의 변화가 유동적이기 때문에 유동성에 적응 가능하고 필터링 없이도 사용자 위치가 예측 가능하도록 가우시안 혼합 모델(GMM, Gaussian Mixture Model)^[20]과 토픽 모델링(Topic Modeling)^[21]을 이용하여 실내공간을 기능적인 측면에서 식별하고 이를 기반으로 사용자가 위치한 공간을 예측한다. 얻어진 사용자 위치정보는 스마트 홈, 홈 메디컬과 같은 다양한 응용분야에서 이용할 수 있다. 예를 들어 특히 스마트 홈 분야에서 사용자의 위치 정보와 함께 스마트 미터기로 측정된 에너지 소모 정보를 이용하여 개인별 에너지 소모량을 예측하거나 실시간으로 에너지 소비를 모니터링 하는 등의 확장을 통해 에너지 절약을 유도하는 데에 큰 도움이 될 것이라 예상된다.

본 논문의 구성은 다음과 같다. 2장에서 수집된 와이파이 신호들을 이용해서 독립된 기능을 갖는 공간들을 대표할 수 있는 특징벡터를 추출하는 방법과 그 정보들을 이용해서 실제 사용자의 위치를 예측하는 방법에 대해서 살펴본다. 그리고 3장에서 실제 가정에서 수행한 실험 결과를 분석하고 4장에서 결론을 맺는다.

II. 실내공간 식별 및 사용자 공간 위치 예측

본 절에서는 실내공간 식별 및 사용자 위치 예측에 관해서 살펴본다. 우선 2.1절에서는 사용자 공간 예측 시스템의 구성에 대해 설명하고 2.2절에서는 토픽 모델 구축에 대해서 마지막으로 2.3절에서는 사용자 공간 위치 예측 모듈에 대해서 살펴본다.

2.1 시스템 구성

본 논문은 일반 가정에서 사용자의 스마트폰으로 수집한 와이파이 신호를 기반으로 사용자가 위치하는 공간을 예측하는 사용자 공간 예측 시스템을 제안한다. 그림 1을 보면 사용자 공간 예측 시스템의 구성이 나타나 있다.

시스템은 크게 토픽 모델 구축 모듈과 사용자 공간 위치 예측 모듈로 구성되어 있다. 토픽 모델 구축 모듈은 특정 공간을 지칭하는 라벨을 갖는 와이파이 핑거프린트 데이터를 이용하여 공간별 특징을 추출하여 공간을 구분하는 역할을 수행하고 사용자 공간 위치 예측 모듈은 실제 와이파이 핑거프린트 데이터로 사용자가 어느 공간에 위치하고 있는지를 판단하는 역할을 수행한다. 토픽 모델 구축 모듈의 데이터 라벨링 단계에서는 와이파이 핑거프린트 데이터에 그 데이터

가 수집된 공간의 식별번호를 덧붙인다. 우선 수집한 와이파이 신호 데이터에 가우시안 혼합 모델을 적용하여 구한 각 가우시안 모델들이 접하는 지점을 기준으로 정한다. 그 뒤에 수집한 데이터를 이용하여 와이파이 핑거프린트를 구성하고 앞서 구한 기준을 참고하여 일정한 시간 단위로 수집된 와이파이 핑거프린트의 집합 별로 토픽 모델링에 필요한 하나의 도큐먼트를 생성하고 생성된 모든 도큐먼트를 사용하여 토픽 모델을 구축한다. 사용자 공간 위치 예측 모듈에서는 실제 사용자의 스마트폰으로 수집한 데이터를 가공하여 토픽 모델 구축 모듈에서처럼 도큐먼트를 생성한 뒤에 구축된 토픽 모델에 입력하여 수집된 데이터가 각각의 공간에 속할 확률 값을 얻는다. 마지막으로 그 확률 값을 확인하여 가장 큰 값을 갖는 공간을 사용자가 위치하는 공간으로 예측한다.

2.2 토픽 모델 구축 모듈

2.2.1 와이파이 핑거프린트 구성과 가우시안 혼합 모델을 이용한 기준 정하기

가우시안 혼합 모델을 이용하여 토픽 모델링에 사용될 기준을 구하기 이전에 수집된 데이터로 벡터를 구성하는 방법에 대해서 살펴보면 다음과 같다. 우선 수집된 데이터 F_t 를 식(1)과 같이 나타낼 수 있다.

$$t \geq 1 \text{ 이고 } m \geq 2 \text{ 일 때,}$$

$$F_t = \left\{ \begin{matrix} ap_1 : r_{ss_1}, ap_2 : r_{ss_2}, \dots \\ \dots, ap_m : r_{ss_m}, Label, Document ID \end{matrix} \right\} \quad (1)$$

여기서 t 는 와이파이 핑거프린트 데이터인 F_t 가 수집된 시간 즉, 타임스탬프를 나타내고 ap_m 과 r_{ss_m} 은 각각 t 시점에 수집된 m 번 째 AP의 물리주소와 신호 세기를 나타낸다. 이 때, $m \geq i \geq 1$ 인 경우 해당하는 와이파이 핑거프린트 F_t 에 속하는 AP의 개수인 m 은 타임스탬프에 따라 변화한다. 또한 $Label$ 은 벡터 F_t 가 수집된 공간의 식별번호를 나타내는데 공간 식별번호의 개수는 한 가정에서 구분하고자 하는 공간의 개수만큼 존재한다. 공간 식별번호은 토픽 모델 구축 모듈에서 사용되며 사용자 공간 위치 예측 모듈에서는 사용되지 않는 정보이다. $Document ID$ 는 수집단위 식별번호로 단위시간 동안에 수집되는 모든 와이파이 핑거프린트는 같은 값을 가지며 2.2.2절에서 도큐먼트를 구성하는 단위로 사용된다.

다음으로 수집한 와이파이 핑거프린트의 신호세기를 적당한 그룹으로 나눌 기준을 선택하는 방법에 대

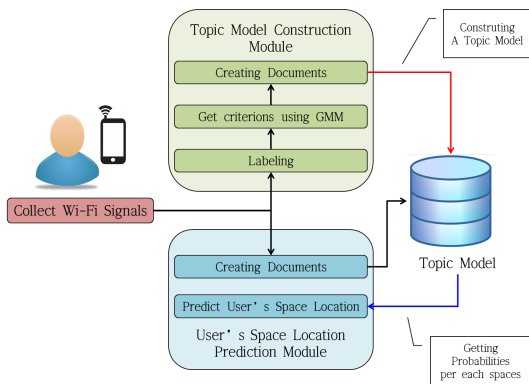


그림 1. 사용자 공간 예측 시스템
Fig. 1. User's Space Location Predicting System

해서 설명한다. 구해진 기준은 후에 공간별로 와이파이 핑거프린트의 특징을 추출하기 위해서 사용하는 토픽 모델링 알고리즘의 입력 데이터로 사용될 도큐먼트를 구성하는 문자열을 생성하는데 사용한다. 그림 2와 그림 3은 한 가정에서 수집된 와이파이 신호 데이터를 각각 다른 방식으로 표현한 그래프이다. 그림 2의 경우에는 단순히 데이터의 평균과 표준편차를 구하여 하나의 가우시안 분포로 나타낸 것이다. 해당 가정에서 수집된 와이파이 신호세기의 전체적 특징을 알 수 있지만 모든 데이터를 하나의 분포로 나타내기 때문에 표준 편차가 크고 신호세기에 따른 특징을 알 수 없다. 반면에 그림 3은 똑같은 데이터에 가우시안 혼합 모델을 적용하여 얻어진 여러 개의 가우시안 분포를 그래프로 나타낸 것이다. 그림 2에서 사용한 것과 같은 데이터를 하나의 분포가 아닌 여러 개의 분포(클러스터)로 나눠주고 있기 때문에 해당 가정에서 수집되는 와이파이 데이터 신호세기의 특징을 더 자세하게 표현할 수 있다. 즉, 한 가정에서 수집되는 두 개 AP의 신호세기가 다르다 가정할 때 그림 2에서는 신

호세기에 따른 두 AP의 차이를 표현할 수가 없지만 그림 3에서는 신호세기에 따라 두 AP가 다른 클러스터에 속한다고 표현 할 수 있다.

본 논문에서는 널리 사용되고 있는 데이터 마이닝 툴인 WEKA의 EM(Expectation-Maximization)알고리즘¹⁷⁾을 사용하였다. EM 알고리즘은 가우시안 혼합 모델 알고리즘을 기반으로 하는 클러스터링 알고리즘이다. 3.2절에서 구한 모든 와이파이 핑거프린트들의 신호세기 데이터에 EM 알고리즘을 적용하면 신호세기에 따른 클러스터들과 각 클러스터에 속하는 신호세기 데이터들의 평균(식 2.1)과 표준편차(식 2.2)를 얻을 수 있다.

$$M = \{m_1, m_2, \dots, m_i, \dots, m_n\} \quad (2.1)$$

$$D = \{\sigma_1, \sigma_2, \dots, \sigma_i, \dots, \sigma_n\} \quad (2.2)$$

위의 식에서 n 은 클러스터의 개수이며 $n \geq 1$ 이다. 이 때 n 의 값은 EM 알고리즘에 의해 자동으로 결정된다. 위의 식에서 M 은 클러스터의 평균들의 집합을 나타내고 D 는 해당 클러스터들의 표준편차 값을 나타낸다. $1 \leq i \leq n$ 일 때 M 의 원소 m_i 와 그에 해당하는 D 의 원소 σ_i 는 각각 i 번째 클러스터에 해당하는 평균과 표준편차를 나타낸다. 각 클러스터의 평균과 표준편차를 가우스 분포 식에 대입해 보면 아래의 식(3)과 같이 나타낼 수 있고 정규분포 그래프로 나타내면 그림 3과 같다.

$$f_i(x) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - m_i}{\sigma_i}\right)^2} \quad (1 \leq i \leq n) \quad (3)$$

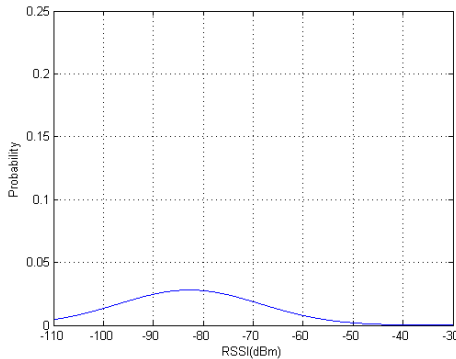


그림 2. 전체 수집 데이터에 대한 하나의 가우시안 분포 Fig. 2. A Gaussian Distribution for Total RSSI Data

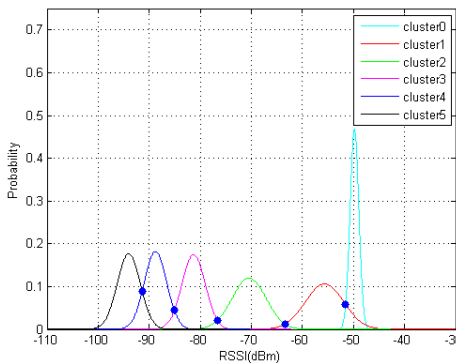


그림 3. EM알고리즘 적용 후 얻어진 6개 가우시안 분포 Fig. 3. Six Gaussian Distributions after Applying EM

본 절에서는 수집한 와이파이신호 데이터를 수신 신호세기 값에 따라 여러 구간들로 잘 구분할 수 있는 기준을 찾아야 하는데 그러기 위해서 그림 3에서 여러 정규분포 그래프가 겹치는 교점(그림에서 파란 점으로 표시된 부분)을 기준으로 구간을 구분한다.

교점을 구하기 위해서 우선 구해진 평균과 표준편차를 평균의 내림차순으로 정렬한다. 결과는 각각 식(4.1), 식(4.2)와 같다. 식(4.1)에서 m'_i 와 식(4.2)에서 σ'_i 는 식(2.1)과 식(2.2)에서와 같이 클러스터 i 에 해당하는 평균과 표준편차이다.

$$M_{ordered} = \{m'_1, m'_2, \dots, m'_i, \dots, m'_n\} \quad (4.1)$$

$$D_{ordered} = \{\sigma'_1, \sigma'_2, \dots, \sigma'_i, \dots, \sigma'_n\} \quad (4.2)$$

$n \geq 1$ 이고 $1 \leq i \leq n$ 을 만족할 때 i 번째 클러스터의 평균인 m'_i 는 그 앞 원소인 $i-1$ 번째 클러스터의 평균 $m'_{(i-1)}$ 보다 작고 다음 원소인 $i+1$ 번째 클러스터의 평균 $m'_{(i+1)}$ 보다 크다. 평균의 내림차순으로 정렬한 뒤 제일 첫 번째 클러스터와 두 번째 클러스터 그 다음에는 두 번째 클러스터와 세 번째 클러스터처럼 앞부터 차례대로 한 쌍의 클러스터를 선택하여 각각의 평균과 표준편차를 가지고 두 개의 가우스 분포식을 구한 뒤에 식(5)와 같은 방정식을 만들어 해를 구한다. 여기서 해들 중 알맞은 하나의 값을 교점으로 정한다.

$$f_{(i-1)}(x) - f_i(x) = 0 \quad (2 \leq i \leq n) \quad (5)$$

위의 식(5)를 만족하는 교점이 여러 개 존재한다면 그 값들 중 적당한 값 하나를 기준으로 선택한다. 기준으로 고려하고 있는 값이 실제 수신 신호세기의 범위인 -100에서 0사이의 값을 만족하고 현 단계에서 살펴보고 있는 값이 바로 이전 단계에서 선택된 기준값보다 작은 값이라는 조건을 만족 한다면 그 값을 기준으로 선택할 수 있다. 이 과정을 통해 결과적으로 총 $n-1$ 개의 기준이 선택되고 기준은 2.2.2절에서 설명하게 될 토픽 모델링에서 사용되는 도큐먼트를 구성하는 문자열을 생성하는데 쓰이게 된다.

2.2.2. 토픽 모델링을 통한 실내공간 식별

본 논문에서는 수집한 데이터를 토픽 모델링 알고리즘을 이용하여 공간 별로 특징을 추출하였다. 토픽 모델링은 주로 기계학습이나 자연언어처리 분야에서 사용되는 알고리즘이다. 토픽 모델은 문서들의 집합 안에서 그 문서를 구성하는 많은 문자열들 중 문서를 대표하는 “토픽”을 문자열들의 빈도를 이용하여 정의하고 그 “토픽”을 문서의 특징으로써 사용하는 통계학 적인 모델이다.

토픽 모델링을 사용자의 위치를 예측하는 시스템에 적용하기 위해 본 논문에서는 1회 수집한 와이파이 핑거프린트들의 집합을 하나의 문서로 정의 하였다. 또한 해당하는 와이파이 핑거프린트 집합 안의 AP 하나하나의 물리주소, 수신 신호세기 정보로 문자열을 생성하여 사용하였다. 가장 먼저 토픽 모델링에 쓰일 도큐먼트를 생성해야 하는데 하나의 도큐먼트는 *Document ID*, *Label* 그리고 많은 *term*들로 구성된다. *Document ID*는 수집된 단위별 식별번호 역할을 하고 *Label*은 그 도큐먼트가 실제로 속하는 공간의

식별번호를 나타낸다. 그리고 *term*은 그 도큐먼트가 갖는 실제 데이터의 문자열을 의미한다. 이를 도식화 하여 나타낸 것이 아래의 그림 4이다.

*Document ID*는 2.2.1절 식(1)에서 와이파이 핑거프린트 F_i 의 *Document ID*에 해당하며 모든 도큐먼트들은 각자 구별되는 도큐먼트 식별번호를 갖는다. 그리고 *Label*은 그 도큐먼트를 구성하는 데이터가 수집된 실제 공간의 식별번호를 의미한다. 그러므로 한 단위시간에 수집된 모든 와이파이 핑거프린트는 하나의 도큐먼트를 구성하는데 사용되며 같은 공간에서 수집된 데이터로 생성한 도큐먼트들은 같은 공간 식별번호를 갖게 된다. 하나의 문자열은 한 개의 와이파이 신호 데이터로 생성하는데 해당 데이터의 물리주소 정보와 신호세기 수신 신호세기의 조합으로 구성된다. 이 때 신호세기는 2.2.가.절에 구한 기준으로 구분한 레벨을 이용한다. 와이파이 핑거프린트인 F_i 의 한 원소(하나의 AP신호 데이터) 당 하나의 문자열이 생성되고 한 공간에서는 비슷한 경향으로 와이파이 신호가 수신되기 때문에 하나의 도큐먼트에는 똑같은 문자열들이 여러 개 존재 할 수 있다. 덧붙여 말하면 어떤 고유한 문자열은 다양한 공간 식별번호를 갖는 도큐먼트들에 걸쳐서 나타날 수도 있고 혹은 같은 공간 식별번호를 갖는 도큐먼트에만 존재를 할 수도 있다. 후자의 경우가 의미하는 바는 해당 고유한 문자열이 그 도큐먼트에서만 다수 나타난다면 그 문자열에 의해서 공간의 특징이 정해질 수 있다는 점이다. 반대로 전자의 경우에는 한 공간의 특징을 추출하는데 있어 그 문자열이 큰 의미가 없다는 것을 의미한다.

수집한 데이터로 도큐먼트의 문자열을 구성하는 자세한 방법에 대해서 살펴보자. 하나의 문자열을 구성하는 데에는 한 개의 와이파이신호 데이터에 대한 물

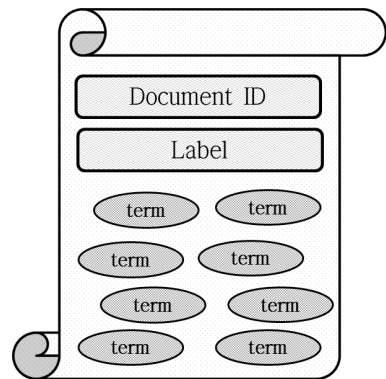


그림 4 도큐먼트 구조
Fig. 4 Document Structure

리주소와 신호세기에 따른 레벨이 이용된다. 신호세기는 일정한 기준에 따라 여러 레벨로 구분된다. 신호의 세기를 몇 개의 레벨로 구분할 것인지, 즉, 여러 레벨로 구분하기 위한 기준이 몇 개인지는 가우시안 혼합 모델 알고리즘을 이용하기 때문에 자동으로 정해진다. 문자열구성에 필요한 신호 세기의 레벨이 정해지면 물리주소와 함께 두 값을 단순하게 연결시키는 것으로 문자열의 생성이 끝난다. 예를 들어 위의 과정을 간단하게 설명하기 위해 2.2.가. 절의 식(1)과 같은 형태로 데이터가 주어졌다고 가정하자. 와이파이 핑거프린트인 $F_i (1 \leq i \leq n)$ 의 한 원소 $ap_j : rss_j (1 \leq j \leq m)$ 가 있을 때 ap_j 의 값은 00:26:66:ff:83:04, rss_j 는 -56이라고 가정하고 이 원소를 도큐먼트의 문자열로 바꿔보자. 우선 ap_j 는 j 번째 AP의 물리주소를 의미하므로 그대로 두고 신호세기를 나타내는 rss_j 를 기준에 따른 레벨로 바꾼다. 레벨을 구분하는 기준이 C1과 C2이고 각각 -45와 -80이라고 가정하면 나올 수 있는 레벨은 $L1(-45 \leq L1 \leq 0)$, $L2(-80 \leq L2 < 45)$ 그리고 $L3(-100 \leq L3 < 80)$ 가 된다. rss_j 에 해당하는 -56이 $-80 < 56 < 45$ 를 만족하므로 rss_j 가 속하는 레벨은 L2가 된다. 그러므로 F_i 의 j 번째 원소인 $ap_j : rss_j$ 의 문자열은 물리주소와 신호세기 레벨을 합한 002666ff8304L2가 된다.

앞서 설명한 과정을 거쳐 얻어진 문자열들로 구성된 도큐먼트들에 토픽 모델링을 적용하였다. 이를 위해 다양한 옵션으로 토픽 모델링 알고리즘을 실행할 수 있는 환경을 제공하는 Stanford Topic Modeling Toolbox^[18]를 이용하였다. 생성된 도큐먼트들은 각 공간의 라벨 정보를 가지고 있기 때문에 토픽 모델링의 세 가지 모델 중에서 라벨이 있는 도큐먼트들을 다루는 방식에 해당하는 LLDA(Labeled Latent Dirichlet Allocation)을 적용하였다. 생성된 도큐먼트에 토픽 모델링 알고리즘을 적용하면 라벨들의 특성을 포함한 하나의 토픽 모델이 생성이 된다. 이렇게 생성된 토픽 모델을 사용자의 위치를 예측하는데 사용하게 된다.

2.3 사용자 공간 위치 예측 모듈

사용자의 위치 예측모듈에서의 과정은 공간구분 모듈의 과정과 매우 유사하다. 우선 사용자의 스마트폰으로 와이파이 신호를 수집한 뒤 일정한 시간 단위로 나눠준다. 그 후 데이터를 가지고 2.2.1절의 가우시안 혼합 모델 알고리즘을 이용하여 구한 기준을 이용해 토픽 모델링에 필요한 도큐먼트를 생성한다. 여기까지

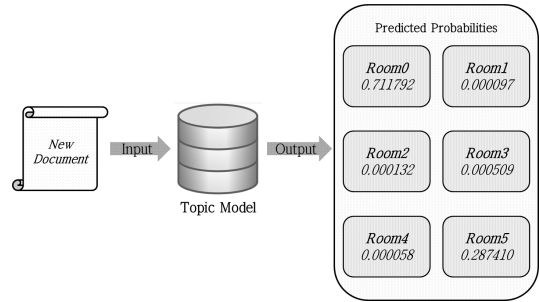


그림 5. 사용자 위치 공간 예측 예제
Fig. 5. An Example of Predicting User's Space Location

는 공간구분 모듈과 흡사하다. 마지막으로 토픽 모델 구축 모듈을 통해서 이미 구축되어 있는 토픽 모델에 새로 생성된 도큐먼트를 입력한다. 그 결과 해당 도큐먼트와 각 공간이 가진 특징들 사이의 유사도 즉, 각 공간에 속할 확률이 나온다. 최종적으로 사용자의 위치를 예측하는 것은 단순히 어떤 도큐먼트가 속할 확률이 제일 큰 공간을 해당 도큐먼트가 속한 공간이라고 결정하는 것과 같다.

예를 들어 위의 그림 5는 토픽 모델 구축 모듈의 예제로 6개의 공간에 대해서 토픽 모델이 구축되어 있는 상황에 새롭게 수집된 데이터로 생성한 한 개의 도큐먼트가 어느 공간에 속하는지를 나타낸 그림이다. 새로운 도큐먼트를 기존의 토픽 모델에 입력한 결과 해당 도큐먼트가 각 공간에 속할 확률이 그림 5의 값과 같이 나왔다고 가정하자. 이때 Room0과의 유사도가 약 0.71로 다른 공간과의 유사도 보다 높음을 알 수가 있다. 결과에 따라 새로운 도큐먼트는 Room1에 속한다고 결정할 수 있다. 그러므로 새로운 도큐먼트가 수집되는 동안 사용자가 머물렀던 공간은 Room1에 해당한다.

III. 실험

본 논문에서는 k -묶음 교차검증법^[22]을 이용한 실험을 통하여 본 논문에서 제안한 사용자 공간 예측 시스템의 성능과 타니모토 거리 측정법(Tanimoto Distance)^[23]을 이용한 성능을 비교분석하였다. k -묶음 교차검증법은 실험을 수행할 데이터의 양이 충분히 많지 않은 경우 사용되는 대표적인 검증방법이다. 본 논문에서는 실제로 여러 가정에서 와이파이 핑거프린트 데이터를 수집하여 실험의 데이터로 사용하였기 때문에 데이터의 양이 제한적이고 따라서 실험 시 k -묶음 교차 검증법을 선택하였다. 또한 본 논문에서

제안한 사용자 공간 예측 시스템의 비교 군으로 타니모토 거리 측정법을 선택하였는데 타니모토 거리 측정법은 와이파이 신호를 포함한 라디오 신호를 이용한 실내에서의 사용자 위치 예측에 있어 비교적 정확한 결과를 내는 알고리즘으로 알려져 있다. 3.1절에서는 실험 환경 및 k -묶음 교차검증법을 이용한 실험 방법에 대해서 3.2절에서는 사용자 공간 예측 시스템과 타니모토 거리 측정법을 이용한 실험 결과를 통하여 성능을 분석한다.

3.1 실험환경 및 실험방법

본 논문에서 제안한 방법의 성능을 분석하기 위해 실제 가정환경에서 데이터를 수집하여 실험하였다. 각 실험환경의 특징은 다음의 표 1과 같다.

와이파이 데이터는 삼성 갤럭시 Note2를 기반으로 다양한 일반 가정집 세 곳에서 수집하였다. 실험환경 가정집의 크기는 평균 $100.9m^2$ 이었으며, 데이터 수집은 거실, 화장실, 방 등 독립된 기능을 갖고 사람들이 주로 머무는 공간에서 수행하였다. 대부분의 가정은 집안에 1~2개의 AP 공유기를 보유하고 있었다. 실험에 사용한 AP 신호는 해당 가정집에서 보유한 AP 신호와 다른 가정집 또는 인접한 건물로부터의 AP 신호로 구성된다. 특성상, 외부 AP 공유기는 직접 설치한 것이 아니기 때문에 개수와 신호세기에서 유동적이고 결과적으로 AP들의 고정성을 보장할 수 없다.

그림 6은 실험환경의 평면도를 나타낸다. 평면도에서 빨간색 별(★) 표시는 공유기가 설치된 위치이며 $R_i (1 \leq i \leq n)$ 는 데이터를 수집한 공간을 의미한다. n 은 위의 표 1의 수집공간의 개수와 일치한다. 그림 6에서 실험에 포함되지 않는 공간(혹은 방)은 거주하는 사람들이 사용하지 않거나 실험환경의 특성상 데이터 수집이 용이하지 않은 공간이다. 네 곳의 실험환경의 차이점을 설명하면 세 실험환경의 위치적인 차이를 보면 (a)는 저층에 위치하고 바로 옆에 다른 건물들과 길이 있으며 (b)와 (c)에 해당하는 가정은 약 20층 정도의 고층에 위치하고 다른 건물들과는 (a)의

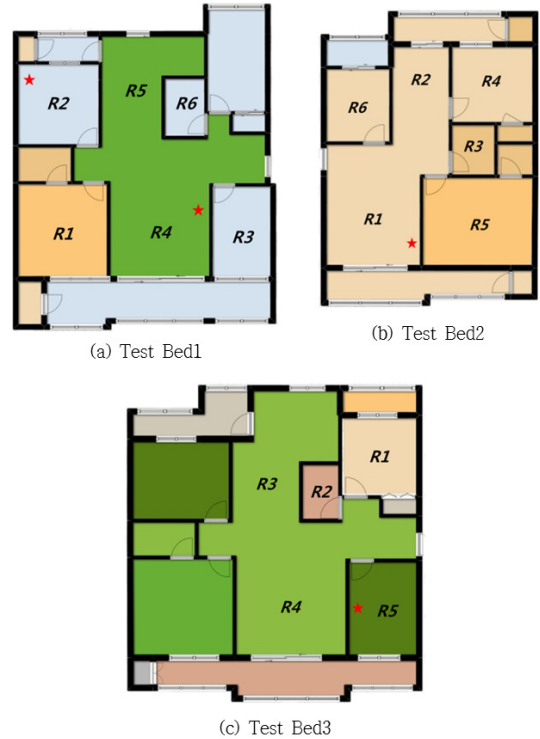


그림 6. 실험환경의 평면도
Fig. 6. The Floor Plans of Test Beds

경우보다 떨어져 있다.

다음으로 와이파이신호 데이터는 실험대상인 각 가정 별로 가족 구성원들이 주로 머무르는 5~6개의 공간을 선택하여 수집하였다. 와이파이 신호 데이터 수집 간격 및 시간은 10초씩 5분로 설정하여 단위시간당 50회 와이파이 신호 데이터를 수집하여 짧은 시간 동안 충분한 데이터 샘플을 얻을 수 있게 하였다. 또한 같은 공간에 대하여 본 논문에서 제안하는 시스템이 일정한 정확도를 보이는지 확인하기 위해 공간마다 와이파이 신호 데이터를 5회씩 수집하였다. 단위시간인 5분 동안 수집된 데이터는 하나의 도큐먼트를 생성하는데 사용되므로 가정에 따라 25~30개의 도큐먼트가 생성되어 각 가정당 25회~30회의 검증을 실시하였다.

본 논문에서 제안하고 있는 알고리즘의 성능을 분석하기 위해 k -묶음 교차검증법을 이용하였다. 실험의 순서는 다음과 같다. 우선 와이파이 신호를 N 개의 공간에 대해서 5회씩 수집하여 생성한 도큐먼트가 $5N$ 개라 하면 그 중에 한 개의 도큐먼트를 제외한 $5N-1$ 개의 도큐먼트들을 토픽 모델링으로 학습시켜 토픽 모델을 구축 한다. 그 뒤 그림 6에서와 같이 학

표 1. 실험환경의 특징
Table 1. The features of Test beds

	Test bed1	Test bed2	Test bed3
Wide	$114.6m^2$	$84.9m^2$	$114.9m^2$
# of Spaces	6	5	5
# of APs	2	1	1
feature	Low-rise apartment house	High-rise apartment house	High-rise apartment house

습을 통해 얻어진 공간별 특징을 갖는 토크모델에 처음에 제외한 한 개의 도큐먼트를 입력하면 입력된 도큐먼트가 N 개의 공간에 속할 확률을 얻을 수 있다. 여기서 해당 도큐먼트가 속할 확률이 가장 큰 공간이 실제 그 도큐먼트가 속한 공간인지 아닌지를 확인함으로써 제안한 알고리즘의 정확도를 얻을 수 있다.

이러한 과정은 모든 도큐먼트를 대상으로 실시하기 때문에 총 $5N$ 회의 교차검증법을 행하게 되고 그 결과로 $5N$ 개의 확률 벡터(공간에 속할 확률들을 모아놓은 벡터)가 나오게 된다. 확률 벡터들을 실제 같은 공간에 속하는 벡터들끼리 묶어 평균을 구하면 최종적으로 총 N 개의 공간에 대한 공간 구분 및 예측 결과를 얻을 수 있다. 그러므로 확률 벡터들의 평균을 확인하여 실제 공간 R_i 에서 수집된 데이터가 본 논문에서 제안한 시스템에 의해 어느 정도의 정확도로 사용자의 위치를 예측하는지 알 수 있다.

3.2 실험결과 분석 및 성능비교

본 절에서는 각 실험환경에서 수집한 와이파이 핑거프린트 샘플 데이터에 대하여 본 논문에서 제안한 사용자 공간 예측 시스템과 타니모토 거리 측정법(Tanimoto Distance) 알고리즘을 비교한다. 그리고 제안한 시스템을 이용한 결과를 실험환경 별로 그래프를 이용하여 자세히 분석한다.

타니모토 거리 측정법은 통계학에서 다양한 샘플들의 유사도를 측정하기 위해서 널리 사용되는 방법이며 RF신호를 이용한 실내에서의 사용자 위치 예측에서도 비교적 높은 정확도를 갖는 방법으로 알려져 있다.

$$f(A,B) = \frac{A \cdot B}{|A|^2 + |B|^2 - A \cdot B} \quad (5)$$

타니모토 거리 측정법의 식은 위의 식(5)와 같다. A 와 B 는 원소의 개수가 같은 벡터이며 f 값은 0에서부터 1사이의 값이 나온다. 이때 f 값이 1에 가까우면 두 벡터가 유사함을 그렇지 않으면 유사하지 않음을 나타낸다. 타니모토 거리 측정법을 사용하여 실내에서의 사용자 위치예측을 위해서는 우선 고정된 AP집합을 가정하여야 한다. 왜냐하면 입력 벡터가 원소의 개수가 같은 벡터이기 때문이다. 본 논문에서는 고정된 AP집합을 가정하기 위해 수집한 데이터를 미리 분석하여 각 가정에서 가장 높은 빈도로 잡히는 10개의 AP를 결정하였다. 원본 데이터에서 고정된 AP집합의 데이터들을 제거한 뒤 벡터에서 잡히지 않는 원소의 값을 최저 신호 값으로 초기화 하고 평균을 구하여

입력벡터를 구성하였다. 여기서 수집 단위는 토크모델링에서 하나의 도큐먼트를 구성하는 단위와 같다. 검증방법은 k -묶음 교차 검증법을 이용하였다. 각 실험당 선택된 하나의 벡터와 각 공간의 특징 벡터들 간의 타니모토 거리를 구한 뒤 가장 1에 가까운 값을 갖는 공간을 선택된 벡터가 속하는 공간으로 예측하였다.

위 그림 7은 그래프는 본 논문에서 제안한 시스템을 이용한 결과와 타니모토 거리알고리즘을 이용한 결과를 비교한 것이다. 가로축은 공간을 나타내며 세로축은 실제로 사용자의 공간 위치를 제대로 예측했는지에 관한 정확도를 나타낸다. 실험환경1에서는 타니모토 거리측정법 알고리즘의 정확도가 조금 더 높다. 그러나 실험환경2에서는 양쪽 모두 정확히 예측을 하고 있고 실험환경3에서는 R3에서 제안한 시스템이 더 정확함을 알 수 있다. 그림 7의 결과로 제안한 시스템과 타니모토 거리측정법 알고리즘이 비슷한 성능을 나타냄을 알 수 있다. 타니모토 거리측정법 알고리즘의 경우에는 고정된 AP집합을 결정하기 위하여 수집한 데이터를 분석하고 그 뒤에 결정된 AP집합에 근거하여 기존의 데이터를 필터링 해야 하는 단계가 필수적이다. 그러나 일반 가정환경은 본인 가정의 AP이 외에 다른 가정의 AP가 고정된 상태라고 가정할 수 없고 따라서 고정된 AP집합을 결정하기 힘든 환경이다. 그러므로 고정된 AP집합 결정 및 필터링 처리가 없이도 비슷한 성능을 보이는 본 논문에서 제안한 사용자 공간 예측 시스템이 타니모토 거리측정 알고리즘보다 가정환경에서 사용자의 실내위치 예측에 유리함을 알 수 있다.

제안한 시스템을 이용한 결과를 실험환경 별로 자세히 분석한 그래프가 그림 8에서 그림 10이다. 그림에서 R_i ($1 \leq i \leq n$, n 대상 공간의 수)은 그림 6의 평면도에 표시되어있는 공간을 의미하고 행은 입력값으로 실제 공간 R_i 의 데이터를 입력했음을 의미한다

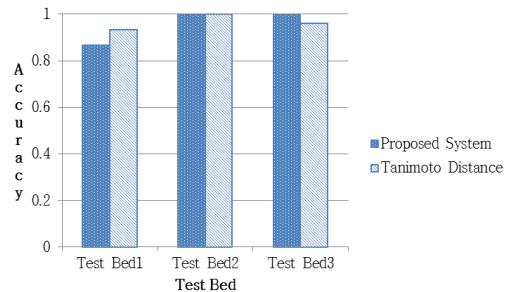


그림 7. 제안한 시스템과 타니모토 거리 측정법의 비교
Fig. 7. The Result compared with Tanimoto Distance

다. 그리고 얻은 예측결과를 의미하며 실제 데이터가 예측 결과에 해당하는 공간에 속할 확률을 의미한다. 예를 들어, 아래의 그림 8은 실험환경1에 대한 실험결과를 나타낸 것이다.

그림 8은 그림 6의 (a)에 대한 결과를 나타낸다. 두 번째 행의 R2에 해당하는 그래프를 보면 실제 데이터가 R2으로 판명 될 확률은 약 67%이며 R5으로 잘못 예측 될 확률은 30%임을 알 수 있다. 3.1.절에서 <표 1>의 실험환경1은 두 개의 공유기를 가지고 있으며 6개의 공간에 대하여 각각 5회씩 와이파이 핑거프린트 샘플을 수집하였다. R2를 제외한 나머지 공간에서 약 75%이상의 정확도로 알맞게 공간을 예측함을 확인할 수 있다.

그림 9와 그림 10은 각각 그림 6의 (b)와 (c)에 대한 결과를 나타낸다. (b)와(c)는 한 개의 공유기를 가지며, (b)에서 6개, (c)에서 5개에 공간에 대하여 샘플을 수집하였다. 각 실험은 공간 수를 제외한 동일한 조건으로 수행하였다. 결과를 보면 (b)와 (c)에서는 평

	R1	R2	R3	R4	R5	R6
R1	0.88	0.00	0.00	0.01	0.00	0.11
R2	0.03	0.67	0.00	0.00	0.30	0.00
R3	0.01	0.00	0.96	0.00	0.00	0.03
R4	0.10	0.02	0.06	0.75	0.06	0.00
R5	0.01	0.11	0.01	0.01	0.78	0.09
R6	0.04	0.00	0.12	0.04	0.03	0.77

그림 8. 실험환경1에서의 결과
Fig. 8. The Result in Test Bed1

	R1	R2	R3	R4	R5	R6
R1	0.98	0.01	0.00	0.01	0.00	0.00
R2	0.00	0.99	0.00	0.00	0.01	0.00
R3	0.00	0.00	0.99	0.00	0.00	0.00
R4	0.01	0.00	0.00	0.99	0.00	0.00
R5	0.00	0.01	0.00	0.00	0.98	0.00
R6	0.00	0.03	0.00	0.00	0.01	0.96

그림 9. 실험환경2에서의 결과
Fig. 9. The Result in Test Bed2

균 95%의 높은 정확도를 보임을 알 수 있다. 두 실험 환경 (b)와 (c)에서의 결과가 (a)에서보다 좋게 나오는 이유는 각 실험환경의 물리적 위치와 관계가 있다. 그림 6의 (a)는 아파트 3층에 위치하며 건물의 바깥쪽에 위치하여 다른 건물과 가깝다. 그에 반해 (b)와 (c)는 각각 아파트 17층과 24층으로 고층에 위치하며 다른 건물과의 거리가 실험환경1의 거리보다 멀다. 즉 (a)은 건물의 바깥쪽에 위치함과 동시에 저층이기 때문에 고층에 위치하는 (b)와 (c)에서 보다 AP 변화의 유동성이 크다. 그러므로 (a)에서의 결과가 (b)와 (c)에서의 결과보다 좋지 않다고 볼 수 있다.

실험초기 단계에서는 일반 가정환경에서 사용자가 위치하는 공간을 예측하는데 가장 영향을 많이 주는 요소가 공간 간의 명확하지 않은 구분이라 예상하였다. 실험환경들은 모두 비슷한 정도로 AP변화의 유동성을 가지고 있다고 가정하였기 때문에 큰 변수가 되지 않을 것이라 예상하였다. 그러나 실험 결과는 실험 환경 인근 다른 건물의 존재 유무나, 인접하는 다른 가정의 수 등을 결정하는 물리적인 위치도 예상과 다르게 AP변화의 유동성에 큰 영향을 끼침을 보여주었다. 이는 실험환경의 물리적인 위치가 사람이나 자동차 등으로 인한 노이즈 발생에 영향을 끼치기 때문이다. 특히 (a)에 해당하는 가정은 저층에 위치하고 인접 건물이 가까운 실험환경의 특성상 AP환경의 변화가 크기 때문에 그림 6 (a)와 같이 다른 실험환경에서의 결과보다 정확도가 낮음을 볼 수 있었다. 즉, 그림 9, 그림 10 그리고 그림 8의 비교를 통해 가정의 물리적 위치에 따른 사용자 위치 공간 예측의 정확도 차이를 볼 수 있었다.

본 논문에서 제안한 사용자 공간 예측 시스템은 일반 가정환경에서 평균 80%이상의 정확도로 사용자의 위치를 예측함을 알 수 있다. 즉, AP 환경의 변동이 많은 일반 가정환경에서 별도의 데이터 필터링이 없

	R1	R2	R3	R4	R5
R1	0.97	0.00	0.01	0.00	0.02
R2	0.00	0.92	0.06	0.01	0.00
R3	0.02	0.00	0.95	0.00	0.03
R4	0.00	0.02	0.00	0.89	0.09
R5	0.00	0.00	0.00	0.01	0.99

그림 10 실험환경3에서의 결과
Fig. 10 The Result in Test Bed3

이도 공간을 구분하고 사용자의 위치를 예측할 수 있는 가능성이 있음을 의미한다. 이는 토픽 모델링 알고리즘의 특성 때문이라고 볼 수 있는데 어떤 공간에서 수집하여 생성한 도큐먼트에 속하는 모든 문자열(AP 물리주소와 신호세기가 속하는 클러스터의 짝)에 대해서 어떤 문자열의 빈도를 그 도큐먼트의 특성으로 사용하기 때문이다. 예를 들어, 한 도큐먼트에 A라는 문자열이 100회, B라는 문자열이 5회 나왔다면 도큐먼트는 빈도가 큰 A 문자열의 영향을 더 많이 받는다.

IV. 결 론

본 논문에서는 넓고 성능이 좋으며 고정된 AP가 다수 존재하는 큰 빌딩환경이 아니라 넓이가 비교적 좁고 소수의 고정된 AP가 존재하는 일반 가정환경에서 사용자가 위치하는 공간을 예측하는 방법을 제안하였다. 이를 위해 첫 번째로 관심 있는 공간에서 수집한 와이파이 수신 신호세기 데이터에 가우스 혼합 모델을 적용하여 신호를 일정한 그룹으로 나눌 수 있는 기준을 얻었다. 그 이후 수집된 데이터들과 기준으로 토픽 모델링에 필요한 도큐먼트를 생성한 뒤 공간 별로 그 공간을 대표하는 Topic Model을 얻었고 마지막으로 사용자의 위치 예측은 앞서 구한 와이파이 핑거프린트에 기초하여 수행되었다.

제안한 시스템을 검증하기 위해 실제 세 곳의 다양한 가정환경에서 수집한 와이파이 데이터로 실험을 수행하였다. 그 결과 평균적으로 약 80%이상의 정확도를 보였다. AP 환경의 변동이 큰 일반 가정환경에서 사용자의 위치를 예측할 수 있는 이유는 수집한 데이터에 대한 특별한 필터링 없이도 공간의 특성을 추출할 수 있기 때문이다. 이를 검증하기 위해 타니모토 거리 측정법 알고리즘과의 비교 결과 필터링 작업이 필요한 타니모토 거리 측정법 알고리즘과 비슷한 성능을 보여 가정환경에서는 제안한 시스템이 더 유리함을 보였다.

향후에는 가정의 물리적인 위치에 따른 AP환경의 유동성을 고려하여 스마트 폰의 가속도 센서, 실내에 설치한 적외선 센서, 초음파 센서의 값 등 다양한 센서의 정보를 사용해 시스템을 확장할 예정이며 본 논문의 시스템을 통해 얻은 사용자 위치정보를 스마트 홈 분야와 연관시켜 개인의 에너지 소비를 모니터링할 수 있는 시스템에 관한 연구를 수행할 계획이다.

References

- [1] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: Indoor location sensing using RFID," *Wirel. Netw.*, vol. 10, no. 6, pp. 701-710, 2004.
- [2] A. Athalye, V. Savic, M. Bolic and P. M. Djuric, "A novel semi-passive RFID system for indoor localization," *Sensors J. IEEE*, vol. 13, no. 2, pp. 528-537, Feb. 2013.
- [3] M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama, "DOLPHIN: A practical approach for implementing a fully distributed indoor ultrasonic positioning system," *UbiComp 2004, Ubiquitous Computing*, pp. 347-365, 2004.
- [4] R. Want, A. Hopper, V. Falcão, and J. Gibbons, "The active badge location system," *ACM Trans. Inf. Syst. (TOIS)*, vol. 10, no. 1, pp. 91-102, 1992.
- [5] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proc. Int. conf. Mobile comput. netw.*, ACM, pp. 32-43, 2000.
- [6] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Antennas Propagation*, vol. 40, no. 2, pp. 207-217, Feb. 1992.
- [7] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. Int. conf. Mobile comput. netw.*, ACM, pp. 173-194, 2010.
- [8] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM 2000*, Vol. 2, pp. 775-784, 2000.
- [9] P. Bahl, V. N. Padmanabhan, and A. Balachandran, *Enhancements to the RADAR user location and tracking system*, Microsoft Research, 2000.
- [10] M. Youssef and A. Agrawala, "The horus WLAN location determination system," in *Proc. Int. conf. Mobile Syst., Appl., Serv.*,

ACM, pp. 205-218, 2005.

[11] M. Azizyan, I. Constandache and R. R. Choudhury, "SurroundSense: Mobile phone localization via ambience Fingerprinting," in *Proc. Int. Conf. Mobile Comput. Netw., ACM*, pp. 261-272, 2009.

[12] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proc. Int. Conf. Mobile Comput. Netw., ACM*, pp. 293-304, 2012

[13] O. B. Kwon and K. S. Kim, "The design and implementation of location information system using wireless fidelity in indoors," *J. Digital Policy & Management*, vol. 11, no. 4, pp. 243-249, Apr. 2013.

[14] H. Jeon, N. Kim, and H. Park, "A study on effective location determination system in indoor environment," *J. KICS*, vol. 34, no. 2, pp. 119-129, Feb. 2009.

[15] J. Oh, "3D indoor positioning system based on smartphone," *J. KICS*, vol. 38C, no. 12, pp. 1126-1133, Dec. 2013.

[16] H. Kim, J. Bae, and J. Choi, "Wireless LAN based indoor positioning using received signal fingerprint and propagation prediction model," *J. KICS*, vol. 38A, no. 12, pp. 1021-1029, Dec. 2013.

[17] Y. Jiang, X. Pan, K. Li, Q. Lv, R. P. Dick, M. Hannigan and L. Shang, "ARIEL: Automatic wi-fi based room fingerprinting for indoor localization," in *ACM Ubicomp*, pp. 441-450, Sept. 2012.

[18] C. Wu, Z. Yang, Y. Liu, and W. Xi, "WILL: Wireless indoor localization without site survey," in *Proc. IEEE INFOCOM*, pp. 64-72, Mar. 2012.

[19] Z. Yang, C. Wu and Y. Liu, "Locating in Fingerprint space: Wireless indoor localization with little human intervention," in *Proc. Int. Conf. Mobile Comput. Netw., ACM*, pp. 269-280, Aug. 2012.

[20] Weka 3: Data mMining Software in Java, Retrieved Dec. 1, 2013, from <http://www.cs.waikato.ac.nz/ml/weka/>

[21] Stanford Topic Modeling Toolbox Retrieved Dec., 1, 2013, from <http://nlp.stanford.edu/software/tmt/tmt-0.4/>

[22] S. Geisser, *Predictive inference: an introduction*, CRC Press, 1993.

[23] T. T. Tanimoto, *An elementary mathematical theory of classification and prediction*, IBM Technical Report, 1958.

안 다 예 (Daye Ahn)



2012년 2월 : 홍익대학교 컴퓨터 공학과 졸업
 2014년 2월 : 홍익대학교 컴퓨터 공학과 석사
 2014년 3월~현재 : 홍익대학교 컴퓨터 공학과 박사과정

<관심분야> 홈 에너지 관리, 스마트 그리드, 스마트 홈, 실시간 시스템 등

하 란 (Rhan Ha)



1987년 2월 : 서울대학교 컴퓨터공학과 졸업
 1994년 2월 : 서울대학교 컴퓨터공학과 석사
 1995년 4월 : University of Illinois at Urbana-Champaign 전산학박사

1989년 3월~1990년 7월 : 한국통신 전임연구원
 1995년 9월~현재 : 홍익대학교 컴퓨터공학과 교수
 <관심분야> 모바일 컴퓨팅, 실시간 시스템, SW보안, 스마트 그리드 등