

# 은닉 마르코프 모델을 이용한 실내 네트워크 맵 매칭

## Indoor Network Map Matching by Hidden Markov Model

김태훈\* · 이기준\*\*

Tae Hoon Kim · Ki-Joune Li

**요약** 최근 다양한 센서들의 성능 개선으로 실내측위가 가능해졌다. 하지만 Wi-Fi 라디오 맵을 이용한 실내 측위나 가속도 센서와 디지털 캠퍼스를 이용한 실내 측위는 아직 상당한 오차를 가지고 있어 지금까지의 연구는 실내 측위의 정확성을 높이는 측위 기술에 대해 많이 진행되었다. 하지만 좌표단위가 아닌 방 단위의 정확성을 가진 실내 맵 매칭이 가능하다면 Wi-Fi 라디오 맵, 가속도 센서 기반의 현재 실내측위기술로도 실내 서비스가 가능하다. 이에 본 연구는 방 단위의 정확성을 가지는 실내 맵 매칭을 위해, 실내 네트워크 맵 매칭에 대해 정의하고, 이를 수행하며 생기는 이슈들에 대해 살펴보고, 이를 해결하기 위해 은닉 마르코프 모델을 사용한 방안에 대해 제시한다.

**키워드** : 실내 네트워크 맵 매칭, 은닉 마르코프 모델(HMM)

**Abstract** Due to recent improvement of various sensor technologies, indoor positioning becomes available. However, Indoor positioning technologies by Wi-Fi radio map and acceleration sensor and digital campus still have a certain level of errors and a number of researches have been done to increase the positioning accuracy of the indoor positioning. If we could provide a room level accuracy, indoor location based services with current indoor positioning methods such as Wi-Fi radio map and acceleration sensors would be possible. In this paper, we propose an indoor map matching method to provide a room level accuracy based on hidden markov model.

**Keywords** : Indoor network based Map Matching, Hidden Markov Model

### 1. 서 론

최근 다양한 센서들의 성능이 개선되고 기기 간의 통신기술이 발달하여 실내에서 위치기반 서비스들이 제공되고 있다. 이때 위치에 따라 반응하는 서비스를 구축하기 위해 필요한 기술은 실내 측위 기술이다. 현재는 스마트폰의 공급이 원활하여 스마트폰에 탑재된 다양한 센서와 WiFi, Beacon 등을 이용하여 추가적인 장비 없이도 과거보다 쉽게 사용자의 실내 측위가 가능해졌다.

하지만 WiFi나 Beacon을 이용한 실내 측위 기술은 오차가 약 5m정도로 실제로 실내에서 사용하기에는 부족함이 있다[8]. 하지만, 실내 공간에서의 위치 기반 서비스는 사람이 위치하는 좌표가 아니라, 방의 정보만 주어져도 서비스가 가능하다. 그러므로 측위 된 좌

표에 대해 좌표수준에서 실내 측위 결과를 보정하기 보다는 방 단위의 정확도를 가지도록 보정해주는 과정이 필요하다. 이러한 보정과정은 맵 매칭을 통해 이루어진다.

맵 매칭은 실세계에서 측위 된 객체 위치와 디지털 맵 위의 네트워크 간에 연결 관계를 만들어주는 과정이다. 이에 실내 맵 매칭은 3차원 좌표를 방의 경계(boundary) 정보를 이용하여 위치를 판별하였을 때 포함되는 방의 식별자를 반환하는 과정으로 볼 수 있다. 이 과정에서는 입력되는 좌표 값이 정확해야 그 결과가 의미가 있다. 하지만 실제 실내에서 측위 된 좌표 값에는 오차가 존재하므로 맵 매칭 시에 입력 값에 대한 오차의 존재를 반영하여 출력 값을 보정해주는 방법론이 필요하다.

본 논문은 다음과 같이 구성된다. 2장에서는 실내

† This research was partly supported by a grant(11 High-tech G11) from Architecture & Urban Development Research Program and by a grant(14NSIP-B080144-01) from National Land Space Information Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government. This work was also partly supported by the ICT R&D program of MSIP/IITP, [14-824-10-005, Development of In-Memory based 3D spatial DBMS for Spatial Convergence Industry] and by BK21PLUS, Creative Human Resource Development Program for IT Convergence.

\* Tae Hoon Kim, Doctor's course, Dept. of Computer Engineering, Pusan National University. Taehoon.kim@pusan.ac.kr

\*\* Ki-joune Li, Professor, Dept. of Computer Engineering, Pusan National University. lik@pnu.edu (Corresponding author)

맵 매칭과 관련된 기존의 연구와 연구 내용을 소개한다. 3장에서는 실내 네트워크 맵 매칭에 대해 정의하고, 실내 네트워크 맵 매칭과 기존의 네트워크 맵 매칭과의 차이점에 대해 설명하고, 실내 네트워크 맵 매칭 시 생기는 이슈에 대해 설명한다. 4장에서는 실내 네트워크 맵 매칭을 하기 위한 은닉 마르코프 모델을 정의한다. 정의한 실내 네트워크 맵 매칭 방법론이 가진 문제점을 해결하기 위해 5장에서는 문제가 생기는 상황에 대해 알아보고, 이를 해결하기 위한 추가적인 방법을 정의한다. 6장에서는 정의한 방법론에 대한 성능 평가를 위한 실험을 하고, 실험의 결과에 대한 설명을 한다. 마지막으로 7장에서는 본 연구에 대해 정리하고 향후 연구를 제시한다.

## 2. 관련 연구

기존의 실내 맵 매칭은 측위 결과를 방 단위의 맵 매칭이 아니고 측위 결과 자체를 실내 지도에 맞춰 보정하는 연구가 대부분이었다.

Xiao는 CRF(Conditional Random Field)를 사용하여 경로를 맵 매칭을 수행하는 알고리즘인 MapCraft를 정의하고 이를 사용하여 기존의 HMM(Hidden Markov Model)을 사용한 방법론에 비해 보다 빠르고 성능이 개선된 결과를 보여주었다[12]. MapCraft는 1단계 맵 전처리, 2단계 상태(state) 및 특징 함수(feature function) 정의, 3단계 특징 가중치 학습, 마지막으로 4단계 시간에 따른 위치 추정의 총 4단계에 걸쳐 수행된다. 하지만 MapCraft는 좌표 수준의 보정을 수행한다. Wu는 WILL(Wireless indoor logical localization)이라는 AP의 위치를 알지 못하더라도 사용할 수 있는 실내 측위 시스템을 개발하였다[11]. WILL은 AP(Access Point)의 위치에 따라 방마다 신호세기가 다른 것과 가속도 센서의 값이 사람의 걸음걸이에 따라 변화하는 것을 이용, 두개의 센서 값을 이용하여 가상 방(virtual room)들과 그 연결 관계를 생성하고 이를 이용해 지문도면(Fingerprint Map)를 구성하여 사람들의 실내 위치를 측위 하는 시스템이다. 하지만 WILL은 방이라는 특성은 고려되었지만 결국 좌표 수준에서의 측위가 목적이며, 그 정확도는 약 80%이다.

측위 결과를 좌표단위로 보정한다는 것에서는 기존의 연구와 동일하지만 방 단위의 정확도를 높이는 연구는 Radaelli가 수행한 연구[7]가 있다. 이것은 복도에 설치된 카메라를 통해 입력되는 비디오 데이터를 영상처리를 이용하여 객체와 방의 입구를 추적하여 객체가 방에 들어가는지 나오는지를 판단하고, 이 정

보를 Wi-Fi 라디오 맵을 이용하여 생성한 측위 데이터와 연결시켜 측위 좌표를 보정해주는 연구가 있다.

또 측위를 좌표가 아닌 기 정의한 의미 있는 단위로 하고 이를 기반으로 위치 추정을 하는 연구는 Kang이 수행한 연구[5]가 있다. 이 연구는 RFID 센서와 태그를 이용하여 위치를 특정하며, RFID가 감지되는 영역과 그렇지 않은 영역을 이용하여 TISS(Trackable Indoor Symbolic Space)를 구성하여 위치를 추정하는 방법에 대해 정의하였다. 결과 태그 감지율이 20%일 때 85% 이상의 위치 추적 성공률을 보여주나, 실제 실험결과가 아닌 시뮬레이션 결과이며 이를 사용하기 위해서는 RFID 태그를 지니고 있어야 하므로 실용성 측면에서 기존의 방법보다 제약점을 지닌다. 이와 비슷한 연구로 특정 위치에 대해 각 AP의 RSS(Radio Signal Strength)의 값들을 사용하여 은닉마르코프모델을 학습시키고, 이를 이용하여 위치를 추정하는 연구를 Kim이 수행하였다[6]. 총 247개의 위치에 대해 101개의 AP에서 발생된 RSS들에 대한 데이터를 학습용과 평가용으로 분리하여 실험을 진행하였는데 결과 학습 시 분산 값에 대해 정확도가 달라지는 것을 확인할 수 있었지만, 제안한 방법을 사용 시 약 54.7%의 낮은 정확도를 가지며 좌표 수준의 위치 추정이 아니라 미리 정의한 위치에 대한 위치 추정이며, 건물의 기하나 실내 네트워크 등의 다른 요소는 고려하지 않았다.

이렇게 방 단위의 실내 맵 매칭에 관한 연구가 없는 이유는 기존의 실내 맵 매칭은 실내 내비게이션이나 로봇의 위치 추적 등이 목적인 연구가 대다수인 것 때문으로 예상된다. 하지만 방 단위의 실내 맵 매칭이 제대로 수행된다면 일정 수준의 실내 LBS(Location Based Service)중 일부 서비스가 가능하다. 예를 들어 집에서 사람이 있는 방을 인식하여 불을 밝혀주거나 TV를 켜주는 등의 스마트 홈 서비스가 가능하며, 백화점에서 고객이 이동하는 점포의 특성에 따라 맞춤형 광고 등이 가능하다. 그러므로 방 단위 정확성을 가진 실내 맵 매칭은 필요하다.

본 논문에서는 방 단위 실내 맵 매칭을 하는 방법으로 실내 네트워크를 이용하였고, 이때 발생하는 이슈를 해결하기 위해 은닉 마르코프 모델을 사용하는 방법을 소개한다.

## 3. 실내 네트워크 맵 매칭

### 3.1 실내 네트워크 맵 매칭 정의

실내 네트워크 맵 매칭이란 객체의 위치 정보를 실

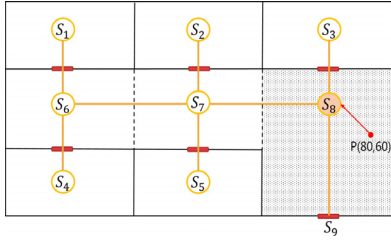


Figure 1. Example of Indoor Network based Map Matching

내 네트워크상으로 가장 적합한 노드와 매치해 주는 과정으로 다음과 같은 함수로 정의 할 수 있다.

$$f((p,t)^*,L) = C; L = (S_c, G_c) \quad (1)$$

여기서 함수의 입력 값 중 하나인  $p$ 는  $(x,y,z)$ 의 형태로 나타낼 수 있는 좌표 값이고, 이를 이용한  $(p,t)^*$ 는 시간에 따른 좌표의 순서열인 궤적이다. 또 다른 입력 값인  $L$ 은 맵 매칭의 대상이 되는 공간으로  $S_c$ 와  $G_c$ 로 이루어져있다.  $S_c$ 는 각 셀의 기하요소 집합이고,  $G_c$ 는 셀 연결 관계 그래프이다. 마지막으로 출력 값  $C$ 는 셀 id를 의미한다.

기본적인 방법으로 방의 경계 정보를 이용하여 객체가 포함된 셀을 찾고, 해당 셀과 매치되는 노드와 매칭 해주는 방법이 있다. 이것을 기본적인 실내 네트워크 맵 매칭(Simple Indoor network Map Matching, SIMM)이라고 한다.

Figure 1은 SIMM의 예시로 측위 좌표  $p(80,60)$ 가 음영 처리된 셀의 내부에 있기 때문에 해당 셀에 포함된  $S_8$ 의 노드에 매칭 되는 것을 확인 할 수 있다.

실내 네트워크를 생성하는 기준 및 방법은 여러 가지 있지만 본 논문에서는 IndoorGML[10]에서 정의된 기준에 맞춰 셀과 네트워크 노드의 관계가 1:1 관계가 되도록 생성한다. 기본적으로 셀마다 노드를 하나씩 생성하는 형태가 되고, 복도와 같은 커다란 셀의 경우에는 하나의 노드로 할 수도, 공간 분할을 통해 셀을 여러 개의 서브 셀로 나눠, 여러 개의 노드를 생성 할 수도 있다. 즉, 셀의 특성에 따라 셀이 나뉘질 수 있다[9]. 하지만, 실내 측위 좌표에는 오차가 존재하여 정확하지 않기 때문에 SIMM만으로는 충분하지 않다.

### 3.2 실내 네트워크 맵 매칭 특징

실내 네트워크 맵 매칭은 실외 네트워크 맵 매칭, 즉 도로 네트워크 맵 매칭과 큰 차이점을 가지고 있다.

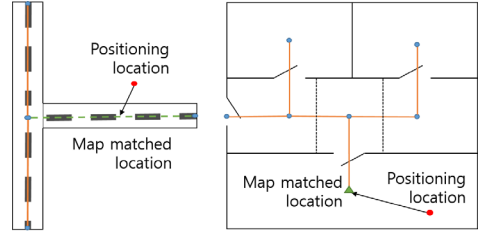


Figure 2. Different Outdoor network map matching from Indoor network map matching

도로 네트워크 맵 매칭은 측위 된 객체의 위치를 도로 네트워크의 엣지에 매칭한다[2]. 하지만 실내 네트워크 맵 매칭에서는 측위 된 객체의 위치를 실내 네트워크의 노드에 매칭한다. 즉, 도로 네트워크에서는 “point to curve” 매칭(혹은 궤적을 사용할 경우 “curve to curve” 매칭)을 하고, 실내 네트워크에서는 “point to polygon” 매칭(혹은 궤적을 사용할 경우 “curve to polygon” 매칭)을 한다. Figure 2은 이러한 차이를 보여주는 예시이다. 원이 측위 된 위치이고 삼각형(혹은 점선)은 매칭 된 결과를 나타낸다.

이러한 차이점이 나타나는 이유는 두 개 네트워크의 특성이 다르기 때문이다. 도로 네트워크에서는 엣지가 도로를, 노드는 도로 간의 연결성을 나타내지만, 실내 네트워크에서는 노드가 셀을, 엣지는 셀간의 연결성을 나타내기 때문이다.

### 3.3 실내 네트워크 맵 매칭 시 이슈

SIMM 시에 발생하는 2가지 이슈에 대해 살펴본다. 먼저 첫 번째 이슈는 SIMM 결과가 네트워크상으로 가능하지 않은 경우이다. 예를 들어 Figure 3의 case 1은 Figure 1의 건물의 실내 네트워크를 이용하여 SIMM 수행 결과가  $S_1 \rightarrow S_6 \rightarrow S_2 \rightarrow S_8$ 와 같이 나온 경우이다. 하지만 이 경우는 네트워크상으로 불가능하다. 이런 경우가 발생하는 여러 가지 상황이 있을 수 있지만 크게 2가지 상황을 통해 일어난 것으로 분류 할 수 있다:

- Case 1-1 : 측위의 샘플링 주기가 늦는 경우 혹은 사람이 뛰어서 이동하는 경우 등 실제로  $S_7$ 을 거쳐  $S_2$ 를 이동한 경우.
- Case 1-2 : 실제로는  $S_2$ 로 이동하지 않고  $S_7$ 으로 이동하였으나, 측위 좌표에 오차가 있어 SIMM의 결과가  $S_2$ 로 나온 경우.

하지만 SIMM의 결과만 가지고는 둘 중 어느 쪽 상황인지 알 수가 없다. 결국 전체적인 경로를 보고 두

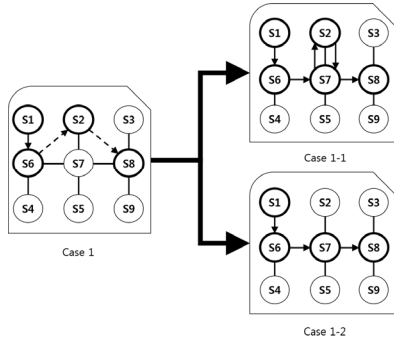


Figure 3. Impossible SIMM results by Indoor Network

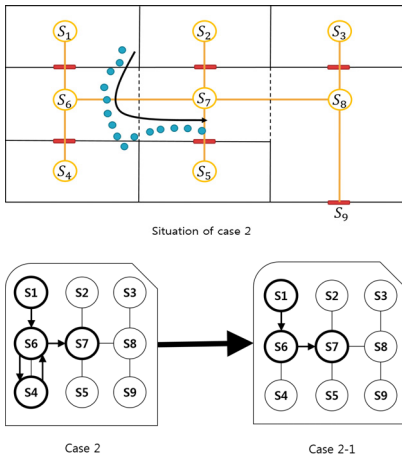


Figure 4. Possible SIMM results by Indoor Network but include error point

가지 상황 중 보다 가능성이 높은 상황으로 SIMM의 결과를 보정해야 한다.

두 번째 이슈는 SIMM 결과에 오차가 존재하지만 네트워크상으로 문제가 없는 경우이다. Figure 4의 case 2의 상황에서 실제 사람은 선을 따라 진행하였을 때, 점과 같이 실내 측위가 되었다고 가정해 보자. 이 경로의 SIMM의 결과는  $s_1 \dots s_1 \rightarrow s_6 \dots s_6 \rightarrow s_4 \rightarrow s_6 \rightarrow s_7 \dots s_7$ 으로 진행을 하는 것으로 나타나 case 2의 결과가 나온다. 이것은 네트워크상으로 가능한 경우이지만 전체적인 경로의 진행상황을 고려하였을 때  $s_4$ 는 오차로 인한 결과일 확률이 큰 것을 짐작 할 수 있다. 그래서 case 2는 case 2-1로 보정을 할 수 있다.

앞의 두 가지 이슈에 대한 결론은 어떤 경우를 선택하기 위해서는 전체적인 경로를 고려하여 보다 확률이 높은 경우를 선택해야 한다는 것이다. 경로에 대한 확률을 구하는 방법으로 본 논문에서는 은닉 마르코프 모델을 사용한다.

## 4. 은닉 마르코프 모델을 이용한 실내 맵 매칭

### 4.1 실내 네트워크 맵 매칭을 위한 은닉 마르코프 모델(Hidden Markov Model)

은닉 마르코프 모델[3]이란 순차적인 데이터에 대해서 현재의 상태는 과거의 상태에만 영향을 받는다는 가정 아래 현재 상태를 추정하는 확률 모델이다. 은닉 마르코프 모델의 요소는 총 5가지로 이뤄진다 : 상태의 개수, 관측 심볼의 개수, 상태 전이 확률분포, 관측 심볼 확률분포, 초기 상태분포 이 중 은닉 마르코프 모델의 객체는 상태와 관측 심볼 2가지다. 여기서 상태는 상태 간의 연결성을 가지고 있고 이러한 연결에는 전이 확률이 있으며, 이를 이용하여 상태는 상태 간의 전이가 가능하다. 그리고 각 상태에는 관측 심볼이 관측될 확률이 존재한다. 이러한 은닉 마르코프 모델의 구조는 실내 네트워크의 구조와 상당히 유사하다. 은닉 마르코프 모델의 상태간의 연결성은 실내 네트워크에서 노드의 연결성과 의미적으로 거의 일치한다. 그래서 은닉 마르코프 모델의 상태를 실내 네트워크의 노드로 정의 할 수 있으며, 이는 1:1로 매칭되는 건물의 셀 식별자와 같다. 그러므로 상태의 개수는 네트워크 노드의 개수와 같다.

관측 심볼 또한 셀 식별자로 설정하였지만, 의미적으로 상태와 차이가 있다. 은닉 마르코프 모델을 사용하면 관측 열에 대해 최고의 확률을 가지는 상태 열을 얻을 수 있다. 그러므로 관측 심볼의 셀 식별자는 입력 데이터로 들어오는 측위 좌표의 SIMM의 결과로 나오는 셀 식별자이고, 상태의 셀 식별자는 이에 대해 가장 최고의 확률을 가지는 보정된 셀 식별자의 의미를 가지게 된다. 이러한 상태와 관측 상태의 의미를 이용하여 은닉 마르코프 모델의 나머지 요소들의 의미를 정의하면 Table 1과 같다.

### 4.2 IHM4: Indoor network based Hidden Markov Model Map Matching

4.1장에서 정의한 은닉 마르코프 모델을 이용한 실내 네트워크 맵 매칭(Indoor network based Hidden Markov Map Matching, IHM4)을 사용하기 위한 요소의 설정법은 다음과 같다.

- **상태의 개수(N):** 각 상태는  $S = \{s_1, s_2, \dots, s_n\}$ 으로 표기한다. 상태는 건물의 셀 식별자를 의미하며, 그 개수는 실내 네트워크의 노드의 개수와 동일하다.

Table 1. Hidden Markov Model for Indoor Network based Map Matching

Element	Meaning
State	Cell ID
Observable Symbol	Cell ID about positioning point, SIMM results
State transition probability	The transition probability between cells
Observation Symbol probability	Probability that the coordinates are included in the cell

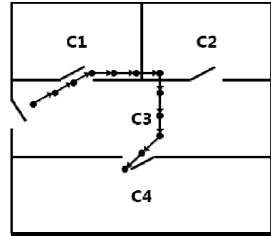
- **관측 심볼의 개수(M):** 각 관측 심볼은  $V = \{v_1, v_2, v_3, \dots, v_n\}$  으로 표기한다. 관측 심볼은 입력된 경로의 SIMM 결과를 의미하며 이는 건물의 셀 식별자와 동일하므로 그 개수는 실내 네트워크의 노드의 개수와 동일하다.
- **초기 상태 분포( $\pi$  matrix):** 어떤 상태가 가장 먼저 나오는지는 입력된 경로의 SIMM 결과의 가장 처음 값과 동일하다. 즉, 처음 시작 위치의 SIMM 결과값의 확률이 1이고 나머지는 0이다. 그래서 시작 측위 좌표의 SIMM 결과는 실제 위치의 실 식별자와 같아야 한다는 가정이 필요하다.
- **상태 전이 확률분포(A matrix):** 상태 전이 확률분포,  $A = \{a_{ij}\}$ 에서  $a_{ij}$ 는 상태  $s_i$ 에서 상태  $s_j$ 로 전이될 확률을 의미하며 다음과 같이 정의된다.

$$a_{ij} = \Pr(q_{t-1} = s_i | q_t = s_j), 1 \leq i, j \leq N \quad (2)$$

상태 전이라는 것은 입력된 경로를 기반으로 발생하기 때문에 입력된 경로의 SIMM 결과의 순서쌍으로 결정한다. 단, 실내 네트워크상으로 가능한 것만을 인정한다. 하지만 기존의 순서쌍만을 대상으로 하면 실내 네트워크상으로 불가능한 경우 보정을 위한 역 경로를 생성할 수 없는 경우가 있으므로 기존의 순서쌍의 순서를 역순으로 바꾼 순서쌍 또한 같이 고려해준다. 제시한 방식으로 상태 전이 확률분포를 설정하면 실내 네트워크상으로 불가능한 결과에 대해 자동적으로 필터가 가능하며 경로 중 있었던 위치에 대해 시간적인 가중치가 함께 고려된다. 예를 들어 Figure 5와 같은 건물에서 표시된 점의 순서로 경로가 주어졌을 때 SIMM의 결과는  $O$ 와 같다. 사실은  $C1$ 과 같이 표현해야 되지만 축약하여 숫자만 표현하였다.

이것을 앞서 정의한 순서쌍을 이용한 방법으로 설정하면  $A$  행렬을 얻을 수 있다. 예를 들어 순서쌍의

$$O = [3, 3, 3, 1, 1, 1, 2, 3, 3, 3, 3, 4]$$



$$A = \begin{pmatrix} 40 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 10 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 0 & 1 & 0 \\ 0.07 & 0.07 & 0.79 & 0.07 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Figure 5. Example: A matrix Setup

숫자가 같은 3,3의 경우에는 행렬의 (3,3)에 2(역 경로의 경우를 포함하여)를 더해준다. 이외의 경우, 예를 들어 순서쌍의 숫자가 1,3인 경우에는 행렬의 (1,3)과 (3,1)에 각각 1을 더해준다 단, (1,2)와 같이 실내 네트워크상으로 불가능한 것은 더하지 않는다.

이렇게 얻은  $A$  행렬을 확률 분포로 사용하기 위해 행렬의 각 행의 합이 1이 되도록 정규화 과정을 진행하면 IHM4에서 사용할 상태 전이 확률 분포를 얻을 수 있다. 각 행의 합이 1이 되어야 하는 이유는 하나의 상태에서 다른(혹은 자신의) 상태로 전이할 수 있는 모든 상황을 고려한 것이기 때문이다.

- **관측 심볼 확률분포(B matrix):** 관측 심볼 확률분포,  $B = \{b_i(v_k)\}$ 는 상태  $s_i$ 에서 관측 심볼  $v_k$ 가 관측될 확률을 의미하며 다음과 같이 정의된다.

$$b_i(v_k) = \Pr(q_t = s_i | q_t = v_k), \leq i \leq N, 1 \leq k \leq M \quad (3)$$

이것은 IHM4에서  $b_i(v_k)$ 는  $s_i$ 에 해당되는 셀에서  $v_k$ 라는 셀이 관측될 확률을 의미한다. 관측될 확률이라는 것은 측위 좌표를 중심으로 오차범위를 고려할 때 해당 좌표가 cell에 속할 확률로 볼 수 있다. 즉, 측위 좌표를 중심으로 일정크기의 원 버퍼를 생성하였을 때, 해당 원 버퍼와 셀의 기하가 겹치는 면적을 측위 좌표가 속하는 상태에서 다른 상태(혹은 같은 상태)가 관측될 확률로 정의할 수 있다. 그러므로 모든 좌표에 대해 각 셀에 대해 겹치는 면적을 확률 매트릭스의 값으로 가산한 다음, 행렬에서 각 행의 합이 1이 되도록 정규화 과정을 거치면 관측 심볼 확률분포를 설정할 수 있다. 행렬의 각 행의 합이 1이 되도록 정규화하는 이유는 하나의 상태에서 관측되는 관측 심볼의 모든 경우를 고려한 것이기 때문이다.

### 4.3 실시간 IHM4

실시간으로 들어오는 측위 좌표에 대해 최적의 셀 식별자를 찾기 위해서는 은닉 마르코프 모델에서 관측 열에 대한 확률 구하는 평가(Evaluation) 문제의 해결법을 적용 할 수 있다. 평가문제의 해결법으로는 Forward 알고리즘과 Backward 알고리즘이 존재하는데 본 논문에서는 Forward 알고리즘을 사용하였다. 실시간 IHM4를 수행하는 단계는 다음과 같다.

**[단계 1]** 주어진 측위 좌표가 관측 될 수 있는 셀의 후보 집합을 생성한다.

**[단계 2]** 만약 해당 측위 좌표가 시작 좌표가 아니라면 이전까지 평가를 통해 가장 높은 확률로 선정된 결과들로 이루어진 경로가 존재한다. 그러면 해당 경로의 마지막에 미리 생성한 후보 집합에서 후보 셀 식별자 선택하여 포함시킨 경로들을 생성한다. 시작 좌표인 경우에는 후보 셀 식별자 하나로 이루어진 경로를 생성한다.

**[단계 3]** 이렇게 생성된 경로들을 Forward 알고리즘을 이용하여 경로에 대한 확률을 구한다.

**[단계 4]** 이 중 가장 높은 확률을 가진 경로를 최종적으로 보정된 경로로 설정한다. 단, 여기서 나온 후보 셀 중 네트워크상으로 가능한 것을 우선적으로 선택하며, 네트워크상으로 가능한 결과가 없는 경우에는 가장 높은 확률을 가진 경로를 선택한다.

이러한 과정을 실시간으로 들어오는 측위 좌표에 대해 반복 수행 하여 실시간 IHM4의 결과를 얻을 수 있다.

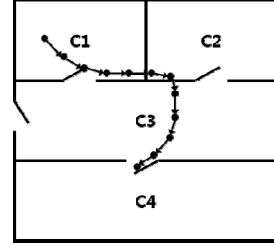


Figure 6. problem occurs when using existing A matrix

우리는 해당 경로에서  $C1 \rightarrow C2$ 가 불가능 한 것을 알고, 이것 대신  $C1 \rightarrow C3$ 로 가는 경로로 보정되기를 기대한다. 하지만  $A_p$ 에서  $C1 \rightarrow C3$ 으로 전이할 확률이 존재하지 않는다. 그러므로  $A_p$ 를 가지고 IHM4를 이용하여 보정된 경로를 생성해도 우리가 기대하는 경로는 생성되지 않는다.

이러한 문제가 발생하는 이유는 주어지는 경로만을 이용하여 상태 전이 확률분포를 생성하기 때문이다. 그래서 이를 보완한 방법은 2가지가 있다.

**- 연결성만을 이용하여 확률분포를 생성하는 방법**

( $A_n$ ): 실내 네트워크에서 노드 간 연결이 있는 경우 1, 없는 경우 0으로 설정한 다음 정규화를 통해 확률 분포를 생성하는 방법이다. 단, 셀프 노드 연결은 무조건 있다. 이 방법을 이용하여 Figure 6에 대해 만든 상태 전이 확률분포  $A_n$ 은 다음과 같다.

$$A_n = \begin{pmatrix} 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

## 5. IHM4의 문제점 및 개선 방법

### 5.1 상태 전이 확률분포 설정 방법

4장에서 제시한 IHM4에 대한 설정 방법론에는 몇몇 특정 상황에 대해 문제점이 존재한다. 먼저 기존의 상태 전이 확률분포 설정 방법의 문제에 대해 살펴본다.

**- 순서쌍을 이용하여 확률분포를 생성하는 방법( $A_p$ ):**

예를 들어 Figure 6와 같은 경로에 대해 기존의 순서쌍을 이용한 상태 전이 확률분포( $A_p$ )의 값은 다음과 같다.

$$A_p = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.67 & 0.33 & 0 \\ 0 & 0.125 & 0.75 & 0.125 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

**- 연결성과 이전의 순서쌍을 이용한 혼합 방법( $A_h$ ):**

먼저 실내 네트워크를 이용하여 확률분포를 한번 설정해 주고, 여기에 기존의 순서쌍을 이용한 방법을 이용하여 전이 정보를 추가적으로 설정하고 마지막으로 정규화를 통해 확률분포를 생성하는 방법이다. 이 방법은 네트워크만 이용하여 만드는 방법에 비해 주어진 경로의 특성을 확률분포에 반영해준다. 이 방법을 이용하여 Figure 6에 대해 만든 상태 전이 확률 분포  $A_h$ 은 다음과 같다.

$$A_h = \begin{pmatrix} 0.9 & 0 & 0.1 & 0 \\ 0 & 0.6 & 0.4 & 0 \\ 0.08 & 0.17 & 0.58 & 0.17 \\ 0 & 0 & 0.67 & 0.33 \end{pmatrix}$$

하지만 실내 네트워크의 연결성을 반영하는 것은

또 다른 문제점을 만든다. 복도와 같이 방에 비해 다른 셀과 많이 연결되는 셀의 경우에 상태 전이 확률분포에서 정규화 과정을 거치면 연결된 상태가 많은 만큼 해당 상태에서 다른 상태로의 전이 확률이 다른 상태에서 해당 상태로의 전이 확률보다 훨씬 작아지게 된다. 예를 들어 복도가 9개의 방과 각각 연결되어 있고 방끼리는 서로 연결되어 있지 않다고 하면 어느 방에서 복도로 이동할 확률은 0.5이지만, 복도에서 어느 방으로 이동할 확률은 0.1이 된다. 이러면 측위 좌표에 대해 다른 요소의 값이 동일한 경우에 복도에서 해당 좌표가 관측될 확률보다 방에서 해당 좌표가 관측될 확률이 더 높아지게 된다.

이를 해결하기 위해서는 각 상태간의 연결 정도를 어느 정도 비슷하게 만들 필요가 있다. 그래서 필요한 작업이 셀 공간분할이다. 본 논문에서는 복도 셀에 대해 연결된 방의 문을 기준으로 공간 분할하는 방법을 사용하였다. 하지만 이러한 공간분할은 관측 심볼 확률분포를 설정하는데 문제가 생기는 역할을 하였다. 이에 대해서는 다음 5.2장에서 설명한다.

### 5.2 관측 심볼 확률분포 설정 방법

기존의 관측 심볼 확률분포 설정 방법은 기존의 설정한 영역의 면적에 겹치는 영역의 면적이 계속 가산된다. 하지만 이런 방법은 어떤 상황에 문제가 발생한다. Figure 7은 셀과 버퍼의 기하 관계에 따른 분류이다.

‘Case 1’은 셀이 버퍼에 포함되는 경우, ‘Case 2’는 셀과 버퍼 사이에 포함관계가 없지만 겹치는 경우, ‘Case 3’은 버퍼가 셀에 포함되는 경우이다.

‘Case 1’인 상황에는 기존의 방법은 문제가 생긴다. 예를 들어 방에 오래 있다가 방 밖으로 나가는 상황을 생각해보자. 이 경우 방에 있는 동안 관측된 좌표에 대해 버퍼의 면적만큼 해당 셀에서 해당 셀이 관측될 확률이 높아진다. 문제는 높아진 만큼 이웃한 셀의 관측 확률이 적어진다. 이런 경우 방 밖으로 이동한 좌표는 방 밖의 셀에서 관측될 확률이 높아진다. 즉, 실제로 측위 좌표의 오차로 방 밖으로 나갔을 수도 있지만, 거기에 대한 고려가 적어지고 SIMM과 유사하게 단순

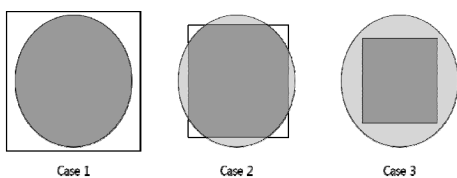


Figure 7. Classification of Cell and Buffer relationship

한 결과가 나온다. 이런 문제점을 해결하기 위해서 겹치는 영역을 가산해 나가지 않고, 특정 셀에 포함되는 측위 좌표의 원 버퍼를 모두 병합하여 하나의 다각형을 만들고 이 다각형이 여러 셀 경계와 겹치는 면적을 특정 셀에서 해당 셀이 관측될 확률로 정의한다. 예를 들어 Figure 8에서 C1에서 C1이 관측될 확률은 C1과 다각형이 겹치는 면적이 되고, C2가 관측될 확률은 C2와 다각형이 겹치는 면적, C3가 관측될 확률은 C3와 다각형이 겹치는 면적이 된다. 물론 C1에서 관측되는 모든 관측 심볼의 확률의 합이 1이 되도록, 각 면적의 비율이 확률 값이 된다.

대부분은 ‘Case 2’의 상황이다. 셀과 버퍼 어느 한쪽에서 포함하는 관계가 아닌 일반적인 관계이기 때문이다.

마지막으로 ‘Case 3’인 상황은 복도를 공간분할을 하였을 때 발생한다. 일반적으로 복도는 방에 비해서 그 폭이 좁고 길게 설계되어 있는 경우가 많다. 이러한 복도 셀을 방의 문을 기준으로 공간분할하면 방에 비해 작은 하부 셀이 생성되는데 이때 버퍼의 크기에 따라 Case 3과 같은 경우가 발생한다. ‘Case 3’이 문제가 되는 이유는 해당되는 하부 셀이 관측될 확률(면적)은 이미 최대가 되었기 때문에 버퍼의 크기가 커지면 상대적으로 그 확률이 작아질 수밖에 없어진다. 이를 보완하기 위해 셀의 면적과 버퍼가 겹치는 면적의 비율을 관측 확률을 설정하였다.

## 6. 실험 결과 및 분석

### 6.1 실험 환경

실험 장소는 부산대학교 자연대 연구 실험동 4층을 대상으로 했다. 대상 건물은 복도를 하나의 셀로 하였을 경우에 총 30개의 셀로 구성되어 있으며, 공간분할을 하였을 경우에는 복도 셀이 총 20개로 분할되어 총 49개의 셀로 구성된다.

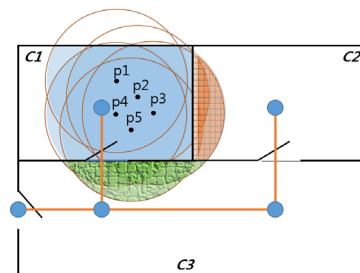


Figure 8. Example for Improved B matrix Setup

실내 측위 데이터 수집은 두 가지로 실시하였다.

- **인공적인 측위좌표 생성:** 프로그램을 이용하여 인공적으로 측위좌표를 생성한 것으로 오류가 정규분포에 따라 생성되도록 오류의 최대 범위만큼의 원 내에 10000개의 임의의 점을 생성하고, 이 중 임의로 5개를 선정, 그 평균을 오류 값으로 이용하였다. 이러한 오류 패턴은 흔히 Wi-Fi등을 이용한 라디오 맵을 이용하여 실내 측위 하는 경우에 나타난다. 현재 센서의 상태를 미리 생성한 라디오 맵과 맞춰 측위하기 때문에 좌표 값이 과거의 영향을 거의 받지 않는다.

- **실제 측위좌표 생성:** 실제 측위좌표는 BuildNGO[1]라는 안드로이드 어플리케이션을 사용하였다. BuildNGO를 사용하기 위해 필요한 대상 건물의 경계 데이터는 JOSM[4]을 이용하여 제작하여 사용하였다. BuildNGO는 초기 위치 측정은 WiFi 라디오 맵을 사용하여 추정하지만, 이후에는 가속도 센서, 자이로 센서, 디지털 컴퍼스 등을 이용하여 측위를 한다. 이때 오류는 현재 측위좌표에 이전에 생긴 오류에 추가로 계속 누적된다.

2가지 생성법에 대해 각각 총 20개의 경로데이터를 생성하였다. 인공적인 측위좌표는 한 경로당 100개의 좌표로 구성되어 총 2000개의 좌표로 구성되며, 실제 측위 좌표는 총 1960개의 좌표로 구성되어 있다. 각 경로 데이터 별로 실시간 IHM4의 결과를 생성하였으며, IHM4의 결과를 생성하는 방법으로는 상태전이 확률분포 설정 시  $A_p, A_n, A_h$ 를 사용한 방법, 관측 심볼 확률분포 설정 시 교차 영역의 가산법, 병합법, 복도의

공간분할 유/무, 셀 면적과 버퍼 면적의 비율을 고려한 방법을 각각 적용하였다. 또한 버퍼의 크기를 0.9m부터 9m까지 0.45m씩 증가시키면서 실제수치와 일치하는 비율을 비교하였다.

### 6.2 인공적인 측위좌표 실험 결과

인공적으로 생성한 실험 데이터는 오차의 반경이 최대 9m가 되도록 오차를 생성하였다. 복도를 하나의 셀로 사용한 경우, 공간분할 한 경우 각각에 대해 총 10개의 경로를 생성하였다. Table 2에 나타난 각 옵션 적용에 따른 실험 결과를 살펴보면, 행렬 A 의 옵션에 따라서는 최적의 원 버퍼 크기가 달라지는 등 결과에 영향을 주지만, 행렬 B 옵션은 결과에 영향을 주지 못하는 것을 확인 할 수 있다. 이러한 결과가 나타난 이유는 생성한 경로가 균등한 속도로 이동하여, 면적을 가산하는 방법이 문제되는 상황이 나타나지 않은 것으로 보인다. 또한 Figure 9와 함께 결과를 보면 복도를 하나의 셀로 사용하는 것과 방 입구를 기준으로 공간 분할한 것의 결과가 유사하나 복도를 하나로 사용하는 것의 결과가 전체적으로 더 나은 것을 확인할 수 있었다. 또한 공간분할을 하였을 경우, 셀의 면적과 버퍼의 면적의 비율을 고려하지 않은 경우에는 버퍼의 크기가 커질수록 SIMM보다 낮게 급격하게 일치율이 떨어지는 것을 확인 할 수 있었다. 또한 버퍼의 크기가 커져감에 따라 일치율이 점차 증가했다가 일정 크기 이후로는 감소하는 경향을 보였으며, 성능이 가장 좋은 구간은 버퍼크기가 3m~6m인 것으로 확인할 수 있었다. 이것은 해당 데이터의 오류가 최대 9m인 정규분포를 따르기 때문에 나타난 경향으로 보인다. 또한 옵션에 따른 성능은 비슷하지만 그중 가장 강건한 경향을 보이는 것은  $A_p$  옵션으로 확인할 수

Table 2. Experimental results of the apply each option (synthetic positioning data)

A matrix option	B matrix option	Another option	Ground truth match rate		Circle buffer radius
			SIMM	IHM4	
$A_p$	Nested =Union	no sub spaced	73.5%	91.9%	4.95m
$A_n$				91.2%	5.85m
$A_h$				91.8%	6.3m
$A_p$	Nested =Union	sub spaced, cell/ buffer ratio	73.5%	89.3%	3.6, 4.95m
$A_n$				89.6%	3.6m
$A_h$				89.5%	4.05m
$A_p$	Nested =Union	sub spaced, no cell/ buffer ratio	73.5%	88.3%	3.15m
$A_n$				87%	2.7, 4.05, 4.5m
$A_h$				87.7%	3.15m

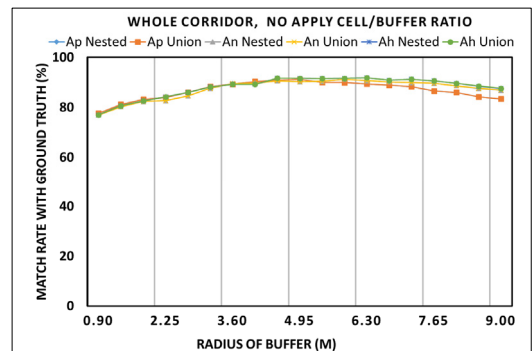


Figure 9. Experimental results of synthetic positioning data by buffer size



Table 3. Experimental results of the apply each option (Real positioning data: SIMM result > 90%)

A matrix option	B matrix option	Another option	Ground truth match rate		Circle buffer radius	
			SIMM	IHM4		
$A_p$	Nested =Union	no sub spaced	98.5%	97.7%	0.9m	
$A_n$				97.4%		0.9, 1.35m
$A_h$				98.2%		
$A_p$	Nested =Union	sub spaced, cell/ buffer ratio	98.5%	97.2%	0.9m	
$A_n$				97.2%		
$A_h$				97.6%		
$A_p$	Nested =Union	sub spaced, no cell/ buffer ratio	98.5%	97.3%	0.9m	
$A_n$				97.0%		
$A_h$				97.8%		

Table 4. Experimental results of the apply each option (Real positioning data : SIMM result < 90%)

A matrix option	B matrix option	Another option	Ground truth match rate		Circle buffer radius
			SIMM	IHM4	
$A_p$	Nested =Union	no sub spaced	76.8%	97.1%	8.55, 9m
$A_n$				94.5%	
$A_h$				94.7%	
$A_p$	Nested =Union	sub spaced, cell/ buffer ratio	76.8%	94.3%	9.45m
$A_n$				94.7%	
$A_h$				94.7%	
$A_p$	Nested =Union	sub spaced, no cell/ buffer ratio	76.8%	89.9%	4.95m
$A_n$				84.1%	
$A_h$				83.7%	

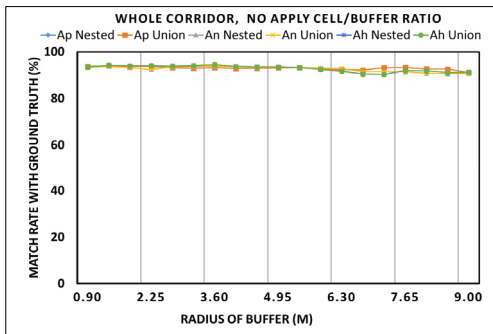


Figure 10. Experimental results of real positioning data by buffer size

있다. 그 이유는  $A_p$  옵션이 다른 옵션에 비해 많은 상황을 고려하지 않고 단순하게 판단하기 때문이라고 예상된다.

### 6.3 실제 측위 좌표 실험 결과

총 20개의 경로를 실제로 이동하며 BuildNGO를 이용하여 측위 데이터를 수집하였다. 실험의 평가는 실제 측위좌표 데이터의 SIMM의 일치율이 90% 이상인 것과 미만인 것을 구분하여 평가 하였다. SIMM의 일치율이 90% 이상인 데이터에 대한 실험 결과는 Table 3과 Figure 10과 같다. BuildNGO의 특성상 이동할수록 오류가 누적되는데, 이러한 일치율이 높은 데이터는 오류가 누적되기 전에 이동이 끝나는 짧은 경로의 경우이다. 결국 오차가 적다는 것은 오차의 반경이 작다는 것이며, 이는 최적의 버퍼 구간이 0.9m~1.35m로

앞쪽에 위치한 것을 통해 확인 할 수 있다. 가장 강건한 경향을 보이는 것은 공간분할을 하지 않은 경우의  $A_h$  옵션이다. 하지만 제한한 방법들의 일치율이 SIMM과 거의 유사하나 약간 낮은 것을 확인 할 수 있다. 이는 셀과 셀 간의 이동시, 즉 문을 통과할 때에 원 버퍼가 한쪽 셀로 치우쳐져야 이동한 것으로 판단하기 때문에 해당 구간에서 약간의 오류가 존재하여 일치율이 떨어지는 것이다.

반대로 SIMM의 일치율이 90%이하인 데이터를 대상으로 제한한 방법론을 적용한 결과는 Table 4와 같다. 결과 매우 높은 일치율을 보여주는 것을 확인 할 수 있다. 일치율이 떨어지는 데이터는 측위 도중 위치가 WiFi 라디오 맵을 통해 보정되어 연속적이지 못하고 갑자기 크게 뛰는 경우 혹은 방에서 나와 복도를 지나가는데 측위 위치가 복도 옆 방 쪽으로 쏠려서 방에서 방으로 이동하는 경우 등이었다. 이렇듯 실내 네트워크의 연결 관계 상 있기 힘든 결과에 대해서 제시한 방법론이 효과적임을 확인 할 수 있었다. 가장 간건한 경향을 보이는 것은 공간분할을 하지 않은 경우의  $A_p$  옵션으로 확인 할 수 있었다. 또한 SIMM의 일치율에 따라 구분한 두 집합에 대한 결과를 종합했을 때 적절한 버퍼의 반경은 발생하는 오차의 반경에 의존적임을 확인 할 수 있었다.

## 7. 결론 및 향후 연구

본 논문에서는 먼저 방 단위의 정확성을 가지는 맵 매칭을 위해 실내 네트워크 맵 매칭에 대해 정의하였

고 이때 발생하는 이슈에 대해 알아보았다. 그리고 이런 이슈를 해결하기 위한 방법으로 확률적인 접근법을 선택하였고, 그 방법으로 은닉 마르코프 모델을 사용하는 방법론을 선택하였다. 이후 실내 네트워크 맵 매칭을 위한 은닉 마르코프 모델을 제시하였으며 이 모델을 사용하기 위한 설정 방법에 대해 정의하였다. 하지만 정의한 설정 방법이 문제가 생기는 특정 상황이 있어 해당 상황을 해결할 수 있는 보완방법을 추가로 제시하였다. 이렇게 제시한 방법론이 생각한대로 맵 매칭 결과를 잘 보정해 주는지 확인하기 위한 실험을 하였다. 실험은 인공적으로 생성한 데이터와 실제 측위 어플리케이션을 이용하여 수집한 데이터를 대상으로 앞서 정의한 여러 가지 방법론(행렬 A, 행렬 B 설정법)에 대해 버퍼의 크기를 변화시키면서 수행하였다.

실험 결과, 인공적으로 생성한 측위 데이터에 대해서는 IHM4의 방법이 실제수치와의 일치율이 최고 91.9%로 SIMM의 일치율인 73.5%과 비교했을 때 개선된 성능을 보여주는 것을 확인하였다. 실제 측위 데이터에 대해서는 SIMM의 일치율이 높은(90%이상)인 경우에는 오히려 일치율이 조금 떨어졌지만, 낮은 경우에는 SIMM의 일치율 76.8%에 대해 최고 97.1%까지 개선되는 것을 확인 할 수 있었다.

또한 원 버퍼의 크기와 오차가 발생하는 범위에 따라 일치율 변화의 경향에 대해 확인 할 수 있었다. 오차가 발생하는 범위가 넓으면 원 버퍼 크기가 증가함에 따라 일치율이 점차 증가하다가 일정 크기 이후로는 감소하는 것을 확인할 수 있었고, 오차가 발생하는 범위가 작으면 원 버퍼 크기가 증가함에 따라 일치율이 감소하는 것을 확인할 수 있었다. 이를 통해 원 버퍼와 오차 발생 범위가 밀접한 관계를 가지고 있음을 확인 할 수 있었다.

향후 연구로는 성능에 크게 영향을 미치는 요소인 버퍼의 크기에 대해 최적의 버퍼 크기를 자동으로 찾기 위한 방안에 대해 알아보고 실제 측위 데이터에서 보이는 오차 패턴인 진행되는 방향에 따라 누적되는 오차에 적합한 방법론에 대해 연구할 것이다.

## References

[1] BuildNGO, 2014, Accessed October 1. <http://www.sailstech.com/#!buildngo-en/c1hnx>.  
 [2] Chung, Y. S; Yoon, H. M; Choi, K. C. 2000, Classification of Map-matching Techniques and A Development, Journal of the Korean society for geospatial information system, 8(1):73-84.

[3] Elliott, R. J; Aggoun, L; Moore, J. B. 1994, Hidden Markov Models: Estimation and Control, Springer, New York.  
 [4] JOSM, 2014, Accessed October 1. <https://josm.openstreetmap.de>.  
 [5] Kang, H. Y; Hwang, J. R; Li, K. J. 2011, Location Tracking in Indoor Symbolic Space with RFID Sensors, Journal of Korea Spaial Information Society, KSIS, 19(3):53-62.  
 [6] Kim, D. H; Hwang, K. B. 2008, Indoor Location Estimation Using Hidden Markov Models with Reliable Variance Estimates, Proceedings of Korea Institute of Information Scientists and Engineers, KISISE, 35(2):311-315.  
 [7] Laura, R; Yael, M; Christian, S. J. 2014, Using Cameras to Improve Wi-Fi Based Indoor Positioning, Paper presented at the annual meeting for Web and Wireless Geographical Information Systems, Springer, May 29-30.  
 [8] Liu, H; Darabi, H; Banerjee, P; Liu, J. 2007, Survey of Wireless Indoor Positioning Techniques and Systems, IEEE Transaction on System, Man, and Cybernetics, Part C, 37(6):1067-1080.  
 [9] Li, K. J; Lee, J. Y. 2013, Basic Concepts of Indoor Spatial Information Candidate Standard IndoorGML and its Applications, Journal of Korea Spaial Information Society, KSIS, 21(3):1-10.  
 [10] OGC IndoorGML, 2014, Accessed October 1, <http://www.opengeospatial.org/projects/groups/indoorgmlswg>.  
 [11] Wu, C; Yang, Z; Liu, Y; Xi, W. 2013, WILL: Wireless Indoor Localization without Site Survey, TPDS Journal, 24(4):839-848.  
 [12] Xiao, Z; Wen, H; Markham, A; Trigoni N. 2014, Lightweight Map Matching for Indoor Localisation Using Conditional Random Fields, Paper presented at the annual meeting for International Conference on Information Processing in Sensor Networks, IEEE Press, April 15-17.

---

Received : 2015.01.09  
 Revised : 2015.03.30  
 Accepted : 2015.04.29