

Improved LTE Fingerprint Positioning Through Clustering-based Repeater Detection and Outlier Removal

Jae Uk Kwon¹, Myeong Seok Chae¹, Seong Yun Cho^{2†}

¹Department of IT Engineering, Kyungil University, Gyeongsan 38428, Korea

²School of Mechanical & Automotive Engineering, Kyungil University, Gyeongsan 38428, Korea

ABSTRACT

In weighted k-nearest neighbor (WkNN)-based Fingerprinting positioning step, a process of comparing the requested positioning signal with signal information for each reference point stored in the fingerprint DB is performed. At this time, the higher the number of matched base station identifiers, the higher the possibility that the terminal exists in the corresponding location, and in fact, an additional weight is added to the location in proportion to the number of matching base stations. On the other hand, if the matching number of base stations is small, the selected candidate reference point has high dependence on the similarity value of the signal. But one problem arises here. The positioning signal can be compared with the repeater signal in the signal information stored on the DB, and the corresponding reference point can be selected as a candidate location. The selected reference point is likely to be an outlier, and if a certain weight is applied to the corresponding location, the error of the estimated location information increases. In order to solve this problem, this paper proposes a WkNN technique including an outlier removal function. To this end, it is first determined whether the repeater signal is included in the DB information of the matched base station. If the reference point for the repeater signal is selected as the candidate position, the reference position corresponding to the outlier is removed based on the clustering technique. The performance of the proposed technique is verified through data acquired in Seocho 1 and 2 dongs in Seoul.

Keywords: fingerprinting positioning, WkNN, clustering, outlier removal

1. INTRODUCTION

최근 위치기반서비스 (LBS: Location Based Service)와 모바일, 사물인터넷, 클라우드 기술 등 여러 정보통신기술 (ICT: Information and Communications Technology)과의 결합에 따라 해당 서비스를 만족시키기 위한 고도화된 기술로 점차 발전하고 있다 (Huang et al. 2018). 이에 따라 위치기반서비스의 요구사항 및 시장의 수요성이 동시에 증가하고 있으며, 서비스의 범주가 대인 추적, 교통정보 등 일상생활에 편의성을 주는 서비스뿐만 아니라 재난 구조와 같은 공공성을 위한 서비스로 응용이 점

차 확대되고 있다.

112/119 호출에 따른 긴급구조 상황에서 요구조자를 성공적으로 구조하기 위해서는 적정 시간 내에 높은 정확도를 갖는 위치 정보를 제공해야 한다. 이를 위해 실외 환경에서는 위성항법장치를 활용한 Global Positioning System (GPS)/Global Navigation Satellite System (GNSS) 정보를 통해 위치를 획득할 수 있다. 그러나 고층 건물이 밀집되어 있는 도심지역이나 실내와 같이 None-line-of-sight (NLOS) 환경을 갖는 음영지역에서는 위성 신호의 차단 및 다중경로의 영향으로 인해 신호가 단절되는 문제가 발생할 수 있다. 이와 반면에 이동통신망 기반 측위 방식은 이미 셀룰러용 기지국 (Base Station)이 전국망으로 많이 설치되어 있기 때문에 실내외 제약없이 모든 공간에서 측위가 가능하다는 장점이 있다. 따라서 본 연구에서는 이동통신망인 Long Term Evolution (LTE)의 수신 전파 (Propagation) 특성 값을 이용하여 Fingerprinting 기법을 기반으로 스마트폰의 위치를 추정하고자 한다.

Fingerprinting 기법은 확률론적 모델링에 의한 측위 방법으

Received Nov 27, 2022 Revised Dec 05, 2022 Accepted Dec 06, 2022

[†]Corresponding Author

E-mail: sycho@kiu.kr

Tel: +82-53-600-5584 Fax: +82-53-600-5599

Jae Uk Kwon <https://orcid.org/0000-0001-6222-5043>

Myeong Seok Chae <https://orcid.org/0000-0002-2638-559X>

Seong Yun Cho <https://orcid.org/0000-0002-4284-2156>

로 노이즈 및 주변 환경 특성을 반영한 실제 Reference Signal Received Power (RSRP) 정보를 위치추정을 위한 정보로 활용하는 방식이다. 이를 위해서는 Fingerprint DB 구축에 필요한 데이터 수집이 이루어져야 한다. 즉, 측위를 위한 공간에서 다수의 샘플 포인트를 설정하고, 각 기준점 (Reference Point)에서 전파되는 RSRP 신호 값과 해당 신호를 송출한 기지국의 Physical Cell Identity 번호와 Band 정보를 획득하도록 한다. 그러나 도심지역과 같이 넓은 공간을 측위 영역으로 설정하는 경우에는 데이터 수집에 대한 시간적 비용이 증가하고, 특정 기준점에서는 지형 및 지리적 문제로 인해 신호를 획득할 수 없는 경우도 발생한다 (Cho 2020). 따라서 LTE 신호 값과 GNSS 위치 정보를 함께 획득할 수 있는 장비를 차량에 탑재 후 도로상을 주행하면서 데이터를 획득하도록 한다 (Cho 2022). 이를 통해 기준점 사이의 공간에서 수집된 위치 기반 RSRP 정보와 공간보간법을 함께 사용하여, 각 기준점상의 신호정보로 변환하도록 한다 (Ismail et al. 2016, Khoo et al. 2022).

위치 추정 단계에서 Fingerprinting 기반으로 단말기의 위치를 계산하는 방법으로는 k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Bayesian Classification, Artificial Neural Network (ANN) 등이 있다 (Xia et al. 2017, Alazab & Gupta 2022). 이들 알고리즘은 특정 데이터에 속한 범위를 예측 및 분류하는 기법으로써 이 중 구현이 간단하면서, 현재 널리 적용되고 있는 기술은 kNN이다 (Xu et al. 2020). 그러나 kNN 기법은 두 가지 문제가 존재한다. 첫 번째는 기준위치가 선택되는 K개의 이웃수에 따라 측위 성능이 각각 다르게 나타나는 것이다. 일반적으로 K의 값은 여러 시행착오를 통해 전체 측위 성능이 가장 높게 추정되는 K 값을 고정하여 사용하게 된다. 그러나 K 값을 고정하여 사용하게 될 경우 특정 측위 지점에서 추정 결과에 대한 성능이 오히려 낮아지는 문제가 발생할 수 있다. 따라서 이 문제를 해결하기 위해 K 값을 적응적으로 적용시킬 수 있는 방법이 제안되었다 (Kwon et al. 2022, Liu et al. 2022). 두 번째는 선택된 기준위치들 중 이상치가 존재하는 경우이다. kNN 기법은 이상치에 대한 정보를 제약없이 그대로 반영하기 때문에 데이터의 노이즈에 민감하다는 단점이 있다. 이로 인해 최종 위치를 계산하는 과정에서 이상치에 해당하는 기준위치를 포함하게 되어 추정된 위치 정보의 전체적인 측위 성능이 낮아지게 된다. 이 문제를 해결하기 위해 본 논문에서는 K-means Clustering 알고리즘을 기반으로 이상치를 제거하는 기법을 제안하고자 한다. 이를 위해 먼저 선택된 후보 기준위치에 대하여 2개의 군집으로 Clustering을 수행하고 두 군집 중심 (Cluster Centroids) 간 거리를 계산한다. 이후 계산된 거리를 이용하여 설정한 임계값 (Threshold)을 기준으로 이상치의 존재 유무를 판단하고, 각 군집에 속한 기준점의 분포도를 파악하여 이상치에 해당하는 기준 위치를 검출하도록 한다. 이 때 제안된 기법을 사용하기 위해서는 DB상에 해당 기지국에 대한 중계기 (Repeater)의 영향이 포함되어 있어야 한다. 중계기의 위치는 기지국에서 송신한 전파가 도달할 수 없는 음영지역에 위치하고 있으며, 기지국으로부터의 전파가 미치지 못하는 영역에 대해 신호를 증폭시켜 다시 전파로 중계해주는 역할을 하고 있다. 이에 따라 데이터를 수집하는 단계에서 중계기의 수신 전파를 갖는 측정치가 함께 획득되며, 해당 신호정보에 대한 특성

이 DB를 생성하는 과정에 영향을 미치게 된다. 즉, 동일한 신호에 대해 선택된 기준점들 중 중계기의 특성이 반영된 기준위치가 이상치로 해당될 가능성이 높아지는 것이다. 따라서 본 연구에서는 측위 과정에서 중계기의 영향으로 인해 선택된 후보 기준위치를 이상치로 판별 후 제거하는 방법에 대해 제안하고자 한다. 제안된 기법에 대한 성능은 서울시 서초 1, 2동에서 획득한 실험 데이터를 기반으로 검증한다.

본 논문은 다음과 같이 구성된다. 먼저 2장에서는 Fingerprinting 기반 측위 알고리즘으로 사용되는 WkNN 기법에 대해 설명하고, 3장에서는 본 논문에서 제안하는 중계기 신호 판단 및 이상치에 해당하는 기준점을 제거하는 방법에 대해 소개한다. 그리고 4장에서는 실험 데이터를 기반으로 제안된 시스템에 대한 측위 성능을 분석하고, 마지막 5장에서 결론을 맺는다.

2. FINGERPRINTING POSITIONING ALGORITHM

측위 단계에서 실제 위치를 추정하기 위해 획득된 LTE 신호 데이터는 DB에 저장된 RSRP와 비교된다. 이 때 사용되는 유사도 거리 공식은 Manhattan 또는 Euclidean 거리를 사용할 수 있으며, 이는 Eq. (1)과 같이 계산될 수 있다.

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

여기서 n 과 m 은 각각 매칭된 기지국과 기준점에 대한 전체 수를 의미하며, s_j 는 측위 단계에서 측정된 j 번째 기지국에 대한 RSRP 값, S_{ij} 는 DB에서 i 번째 기준점에 대한 j 번째 기지국의 RSRP 값을 나타낸다. 그리고 q 의 값이 1 또는 2로 정의될 경우 이는 각각 Manhattan 거리와 Euclidean 거리로 계산될 수 있다.

WkNN 기법은 신호 차이에 대한 유사도 순위가 높은 기준위치에 대해 비교적 큰 가중치를 적용하여 최종적인 위치를 계산하는 방법이다. 다수의 기지국으로부터 송출된 신호 데이터의 조합은 획득 위치마다 각각 다른 조합으로 수집된다. 즉, 셀 기지국의 위치 및 안테나 방향에 따른 송신 커버리지 (Coverage) 반경이 각각 상이하기 때문에 측정치를 획득하는 수집 위치에 따라 송출되는 기지국의 수와 신호의 세기 정보는 전부 다른 특성을 가지게 된다. 이에 따라 매칭되는 기지국의 수가 많은 기준점일수록 해당 위치에 단말기가 존재할 가능성이 높아지게 되고, Eq. (2)와 같이 가중치를 계산하는 과정에서 기지국의 매칭 수에 따라 추가적인 가중치를 부여하도록 한다.

$$\omega_i = \frac{(N_i)^n}{D_i + \lambda} \quad (2)$$

여기서 ω_i 과 D_i 는 각각 i 번째 기준위치에서 계산된 가중치와 신호의 유사도 거리 값을 나타내고, N_i 는 매칭된 기지국 정보의 수, n 은 매칭된 기지국의 수에 추가적인 가중치를 부여하는 제곱 상수항을 나타낸다. 그리고 λ 는 동일한 신호에 대해 분모가 0이 되는 것을 방지하기 위한 일정 상수 값을 의미한다.

Eq. (2)에서 계산된 가중치의 크기를 통해 K개의 후보 기준위

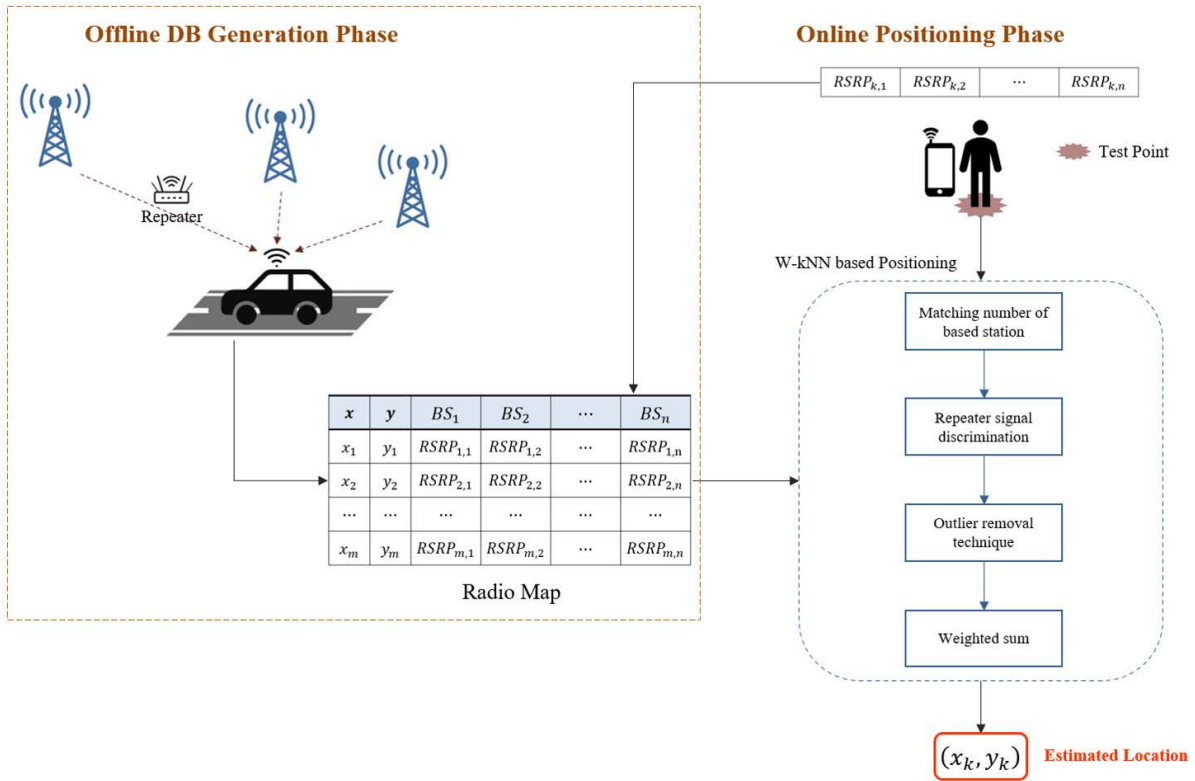


Fig. 1. Architecture for outlier removal system.

치를 선정할 수 있으며, Eq. (3)과 같이 각 가중치를 기준위치의 역수로 할당하여 최종 위치를 계산할 수 있게 된다.

$$(Lat', Lon') = \frac{\sum_{i=1}^K \omega_i (Lat_i, Lon_i)}{\sum_{i=1}^K \omega_i} \quad (3)$$

측위 과정에서 선택되는 기준위치의 정확도는 기지국의 매칭 수에 따라 증가하며, 최종 위치를 계산하기 위한 가중치의 영향도 비례하여 높아지게 된다. 그러나 기지국의 매칭 수가 적은 경우에는 신호의 유사도 값에 의존도가 높아지게 되는데, 이 때 DB 상에서 선정된 후보 기준점들 중 중계기의 영향이 있는 위치가 다수 선택된다면 해당 기준위치는 큰 오차를 포함하게 된다. 이로 인해 단말기의 위치를 계산하는 과정에서 이상치에 해당하는 기준위치에 대해서도 일정 가중치를 적용하기 때문에 최종적으로 출력된 위치정보의 추정 성능이 낮아지는 문제가 발생하게 된다.

3. PROPOSED ALGORITHM

이 장에서는 DB 상에 해당 기지국의 중계기 신호가 저장되어 있는지 존재 여부를 판단하는 기법과 이상치에 해당하는 기준위치를 제거하는 알고리즘에 대해 소개한다.

Fig. 1은 제안된 알고리즘에 대한 전체 시스템의 아키텍처를 나타낸 것이다. 이 시스템은 DB 생성을 위한 데이터 수집 과정을 포

함하여, 실제 측위 과정에서 적용한 제안된 알고리즘의 전체 순서도를 나타낸다. 도심지역에서는 신호 획득을 위해 구분한 기준점의 간격이 클 수밖에 없으며, 특정 공간에서는 측정치를 획득할 수 없는 경우도 존재한다. 따라서 차량에 탑재된 신호 측정 장비를 통해 도로상에서 신호를 획득하도록 하며, 수집데이터와 공간보간법을 함께 사용하여 기준점상의 신호정보로 변환하도록 한다. 이 때 변환하는 방법은 Inverse Distance Weighting 기반의 kNN 기법을 사용한다. 이 기법은 인근의 가까운 측정치로부터 선형으로 결합된 가중치를 통해 새로운 신호정보를 생성할 수 있다는 장점이 있지만, 측정치들과 거리가 먼 곳에 위치하는 기준점의 신호정보를 추정하는 경우 오차가 클 수 있다는 단점이 존재한다. WkNN 기반 측위 단계에서는 새로운 신호 정보가 획득되면 먼저 K개의 후보 기준위치를 선정한다. 이 때 매칭되는 기지국의 수가 적은 경우 DB 상에 저장된 해당 기지국의 수집 데이터에서 중계기의 신호가 포함되어 있는지 판단한다. 이 후 만약 중계기의 신호가 검출되었다고 판단되면 이상치에 해당하는 기준위치를 판별 후 제거하도록 하고, 기지국 신호에 대한 기준점만을 사용하여 단말기의 최종 위치를 계산하도록 한다.

3.1 Repeater Signal Detection Technique

DB 생성을 위해 수집된 데이터에서 기지국별로 중계기 신호의 영향을 판단하는 기법에 대해 설명한다. Fig. 2a는 측정치의 획득 위치를 지도상에 나타낸 것으로 검정색 세모는 PCI-278번 기지국의 위치, 빨간색 점들은 해당 기지국으로부터 송출된 수집데

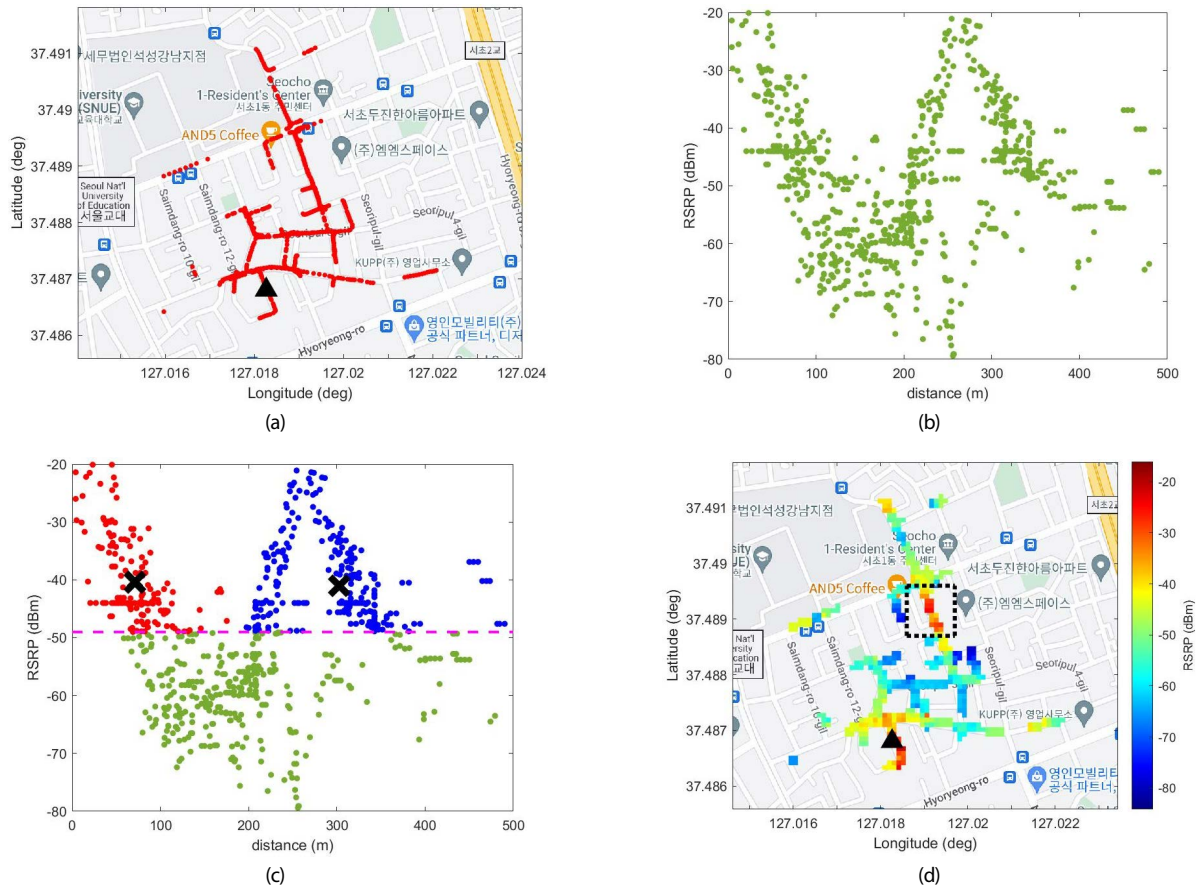


Fig. 2. Repeater signal detection results for PCI-278. (a) location of measurements on the map, (b) RSRP strength for distance, (c) clustering results for collected data, (d) measurements-based DB on the map

이터의 획득 위치를 나타낸 것이다. 해당 데이터를 통해 Eq. (4)의 거리 공식을 이용하여 기지국의 위치를 기준으로 모든 측정치와 의 거리를 각각 계산한다.

$$d_k = \sqrt{dN_k^2 + dE_k^2} \quad (4)$$

$$\begin{aligned} dN_k &= (Lat_{BS} - Lat_{MS-k}) R_m \\ dE_k &= (Lon_{BS} - Lon_{MS-k}) R_t \cos(Lat_{MS-k}) \end{aligned} \quad (5)$$

여기서 $[Lat_{BS} \ Lon_{BS}]^T$ 와 $[Lat_{MS-k} \ Lon_{MS-k}]^T$ 는 각각 기지국과 k 번째 측정치에 대한 수평 위치 벡터이며, 위도와 경도방향의 지구 타원체 반지름은 Eq. (6)과 같이 계산된다.

$$\begin{aligned} R_m &= R_0 (1 - e^2) / \sqrt{(1 - e^2 \sin^2 Lat_{MS-k})^3} \\ R_t &= R_0 / \sqrt{1 - e^2 \sin^2 Lat_{MS-k}} \end{aligned} \quad (6)$$

여기서 $R_0=6,378,177$ [m]와 $e=0.081819$ 는 각각 적도 반지름과 지구 이심률을 의미한다.

Fig. 2b는 Fig. 2a에서 나타난 지도상의 측정치를 모두 사용하여 거리에 따른 RSRP 신호 세기를 나타낸 것이다. 기지국으로부터

터 약 500 m 거리까지 측정치가 획득된 것을 Fig. 2b를 통해 알 수 있으며, 처음에는 거리에 따라 신호 감쇄가 이루어지다가 약 200 m 인근에서부터 신호의 세기가 점차 증가하는 것을 확인할 수 있다.

Fig. 2b에서 나타난 측정치를 이용하여 중계기의 신호를 검출하기 위해 K-means clustering 알고리즘을 수행한다. K-means 알고리즘은 데이터 세트에 포함된 각 요소를 하나의 그룹에만 속하도록 K개 군집으로 분류하는 반복 clustering 기법이다. 이 기법은 상관관계 (correlation) 기반 거리 공식을 분류 척도로 사용한다. 데이터 세트에 대한 K개의 군집 수가 주어지면 거리의 평균 값을 계산하여 초기 중심을 선정한다 (Yuan & Yang 2019). 데이터 세트와 분류할 범주인 K개 군집의 경우 Euclidean 거리를 유사성 지수로 선택하고, Eq. (7)을 통해 데이터 세트와 모든 군집 중심들 간 거리의 제곱합을 계산하여 비용 함수를 최소화하는 것을 목표로 학습한다.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|(x_i - c_j)\|^2 \quad (7)$$

여기서 J 는 비용 함수 (Cost Function), k 와 n 은 각각 군집 중심의 개수와 데이터의 전체 수를 나타내고, x_i 는 데이터 집합의 i 번째 지점, c_j 는 j 번째 군집 중심을 의미한다.

Eq. (7)을 이용하여 업데이트된 군집 중심의 위치를 Eq. (8)과 같이 구할 수 있다.

$$\begin{aligned}\frac{\partial}{\partial c_j} &= \frac{\partial}{\partial c_j} \sum_{i=1}^k \sum_{i=1}^n (x_i - c_j)^2 \\ &= \sum_{i=1}^k \sum_{i=1}^n \frac{\partial}{\partial c_j} (x_i - c_j)^2 \\ &= \sum_{i=1}^n 2(x_i - c_j)\end{aligned}\quad (8)$$

여기서 $\frac{\partial}{\partial c_j}$ 는 j 번째 군집 중심에 대한 편미분을 의미하며, 만약 $\frac{\partial}{\partial c_j}$ 를 0으로 가정한다면 Eq. (9)를 통해 군집 중심의 갱신된 위치를 계산할 수 있게 된다.

$$c_j = \sum_{i=1}^n x_i \quad (9)$$

Eq. (7) 부터 Eq. (9) 까지의 과정을 반복하면서 군집 중심의 위치가 변경되지 않거나 설정된 반복 횟수 (Iteration)에 도달하게 되면 해당 시점에서 출력된 군집 결과를 중계기 신호 검출을 위한 파라미터로 사용하도록 한다.

Fig. 2c는 Fig. 2b의 결과에서 나타난 측정치를 이용하여 두 개의 군집 중심을 기준으로 Clustering한 결과를 나타낸 것이다. 이때 RSRP 신호의 평균 값보다 높은 신호 세기를 갖는 측정치만을 이용하여 Clustering을 수행하였다. 해당 그림에서 나타난 두 군집 중심의 위치를 이용하여 x 축 간 거리가 일정 임계값 이상으로 크게 계산되면 해당 기지국으로부터 수신된 DB용 수집데이터는 중계기 신호를 포함하였다고 판단하였다. 즉, Fig. 2c에서 자주색 파선은 RSRP 평균 신호에 대한 경계선을 의미하고, 검정색 'x'는 두 개의 군집에 대한 중심, 파란색 점과 빨간색 점은 각 군집에 속한 측정치를 각각 나타낸 것이다. Fig. 2d는 측정치 기반으로 생성된 DB를 지도상에 Colormap 형태로 나타낸 것이다. 그림을 보면 기지국에서부터 공간을 전파하면서 신호가 점차 감쇄되는 특성을 보이다가 검정색 점선으로 표시한 영역에서부터 신호가 점차 증폭되는 것을 볼 수 있다. 이를 통해 해당 영역에서는 중계기가 존재할 수 있음을 유추할 수 있게 되고, 제안된 기법을 위해 사용된 Fig. 2c에서도 동일한 거리에 대해 신호의 세기가 점차 증가하는 것을 확인할 수 있다. 따라서 이러한 영향으로 인해 두 군집 중심 간 x 축 거리는 커지게 되며, 해당 기지국으로부터 수신된 DB용 수집데이터는 중계기 신호가 포함되었다고 판단할 수 있게 된다.

3.2 Outlier Removal Technique

WkNN 기반 측위 과정에서 기지국의 매칭 수가 적을 경우에는 신호의 유사도 크기에 따라 기준점들이 선정된다. 즉, 신호 차이에 대한 거리 값이 가까운 순서대로 이웃 수의 기준점들이 선정되는 것이다. 그러나 이웃 수의 후보 기준점에서 중계기의 영향을 갖는 기준점이 다수 포함될 수 있다. 즉, 매칭된 기지국의 DB 정보에서 중계기 신호가 저장된 기준점들이 후보 위치로 선택될 수 있는 것이다. Fig. 3a는 측위 지점에서 획득한 데이터를 이용하여 실제 추정을 위해 선택된 기준점 위치를 나타

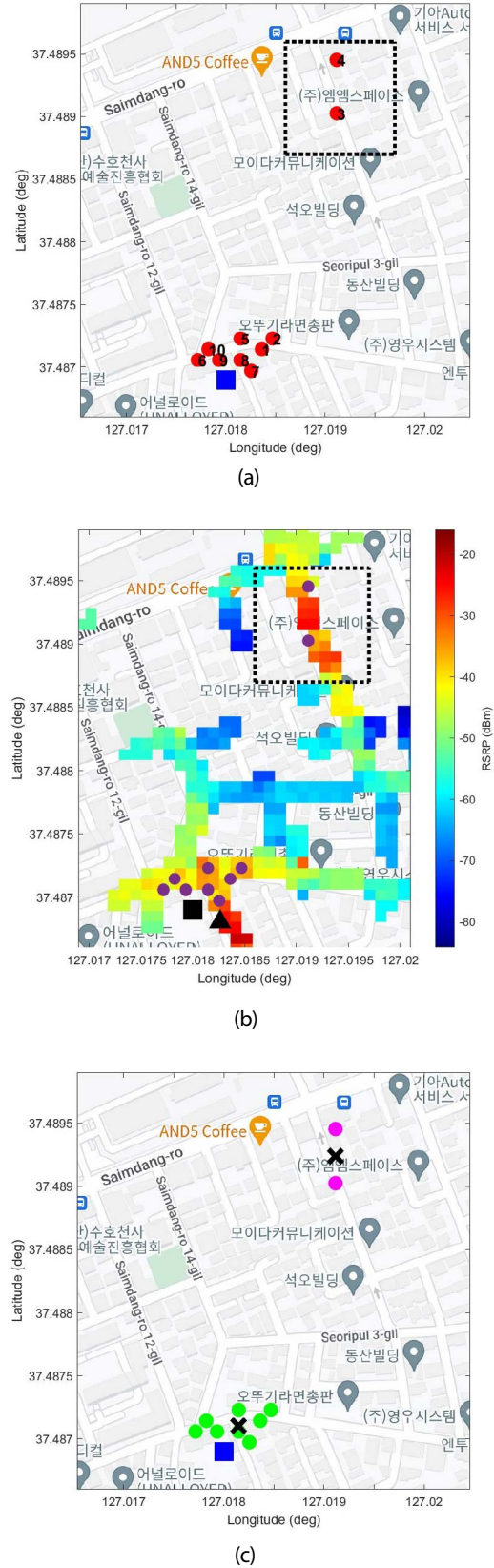


Fig. 3. Outlier remover technique application results. (a) location of the selected candidate reference point, (b) location of reference points on the DB (PCI-278), (c) clustering-based outlier detection

낸 것이다. 여기서 파란색 점은 참 위치, 빨간색 점들은 후보 기준위치를 신호의 유사도 순위에 따라 나타낸 것이다. 그림을 보면 10개의 기준위치가 선택된 것을 볼 수 있으며, 검정색 점선으로 표시한 영역에서 이상치에 해당하는 기준점이 다수 포함된 것을 확인할 수 있다. Fig. 3b는 Fig. 3a에서 나타낸 기준점들을 Fig. 2d의 DB 상에 나타낸 것이다. 여기서 검정색 네모와 세모는 각각 측위 지점과 PCI-278번 기지국의 위치, 보라색 점들은 선택된 후보 기준점들을 의미한다. 대부분의 기준점들은 기지국 신호의 영향이 있는 참 위치 인근으로 추정되는 것을 Fig. 3b를 통해 알 수 있으며, 또한 중계기 신호로 판단되는 DB의 영역에도 후보 기준점이 다수 선택된 것을 확인할 수 있다. 따라서 이상치에 해당하는 중계기 신호에 대한 기준점을 제거하기 위해 다음과 같은 방법을 사용한다. 먼저 선택된 후보 기준점을 Eqs. (7-9)의 과정을 통해 2개의 군집으로 K-means Clustering을 수행하였고, Clustering된 결과를 Fig. 3c와 같이 나타내었다. 이 후 두 군집 중심 간 Euclidean 거리를 계산하고, 해당 거리가 일정 임계값 이상으로 큰 값을 가지게 된다면 분포도가 낮은 군집에 속한 기준위치를 이상치로 판단 후 제거하도록 한다. 즉, Fig. 3c에서 자주색 점에 속한 기준점은 이상치로 판단 후 제거하도록 하였고, 초록색 점에 속한 기준점만을 사용하여 최종적인 위치를 산출하기 위한 파라미터로 사용하였다. 만약 이상적인 상황일 경우 선택된 후보 기준위치는 모두 참 위치 인근에 존재할 것이며, 검정색 'x'로 표시한 군집 중심 간 거리는 낮은 수치로 계산된다. 그러나 중계기의 영향으로 인해 이상치에 해당하는 기준점이 다수 선택되면 군집 중심 간 거리는 큰 값을 가지게 된다. 이 때 이상치 판단 기준이 되는 임계값의 크기는 기지국의 수신 커버리지와 중계기 설치에 따른 단말기의 신호 감도를 고려하여 약 200 m로 설정하였다.

4. EXPERIMENTAL RESULTS

본 논문에서 제안하는 이상치 제거 기능을 포함한 WkNN 기반 측위 기법의 성능을 검증하기 위해 실시할 수 있었다. 시험을 위한 테스트 지역은 서울시 서초 1,2동이며, Galaxy 단말기에 탑재된 Nemo Handy 소프트웨어를 이용하여 측정치를 획득하였다. 차량 시험을 통해 도로상에서만 데이터를 획득하도록 하였으며, 연구과정에서 사용된 기지국의 위치정보는 KT로부터 실험 및 연구목적으로 제공받았다. 먼저 DB용 수집데이터 중에서 측정치의 수가 많고, 신호의 증폭 현상을 구분할 수 있는 기지국 8개를 선정하여 중계기 신호 검출 기법을 적용하였다. Fig. 4는 사용된 기지국의 위치와 DB용 수집데이터의 획득 위치를 지도상에 나타낸 것이다. 해당 그림에서 파란색점들은 측정치를 획득한 위치, 초록색 세모는 기지국의 위치를 PCI 식별자로 표시한 것이다. 차량을 통해 데이터를 수집하므로 측정치의 획득 위치가 도로상에서만 나타나는 것을 확인할 수 있으며, 획득된 데이터의 수는 총 269,212개로 전부 DB를 구축하기 위한 정보로 사용되었다.

Fig. 5는 Fig. 4에서 나타낸 기지국의 위치정보와 DB용 수집데이터 중 기지국별 측정치를 이용하여 중계기 신호를 검출한 결과를 나타낸 것이며, Table 1은 중계기의 영향을 판단하기 위해

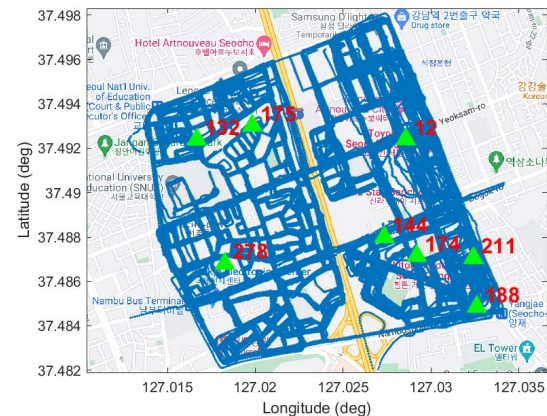


Fig. 4. Location of base stations and measurements.

Table 1. Calculated distance between cluster centers.

PCI	X-axis distance (m)
12	75.2
132	56.1
144	53.7
174	186.0
175	209.7
188	49.1
211	265.7
278	243.7

계산된 군집 중심 간 x축 거리를 나타낸 것이다. 각 그림에서 자주색 파선은 평균 신호의 경계선, 파란색과 빨간색 점들은 각 군집에 속한 측정치, 검정색 'x'는 군집 중심을 나타낸 것이다. 먼저 174, 175, 211, 278번 기지국으로부터 획득된 측정치를 보면 약 200 m의 범위 내에서는 수신 거리가 멀어질수록 신호가 점차 감소되는 경향을 보이다가 일정 거리에서부터 신호의 세기가 증가하는 현상이 동일하게 나타나는 것을 확인할 수 있다. 이를 통해 해당 기지국별 수집데이터는 신호가 증폭되는 구간에 중계기의 영향이 존재할 수 있음을 유추할 수 있으며, 실제로 이러한 영향으로 인해 군집 중심 간 x축 거리는 약 200 m 이상으로 계산되는 것을 확인할 수 있다. 이와 반면에 12, 132, 144, 188번 기지국으로부터 수신된 측정치는 경로 감쇠 모델의 Log-distance의 형태로 수신 거리에 따라 신호 감쇠가 이루어지는 경향을 볼 수 있다. 즉, 일정 수신 거리를 지남에 따라 신호가 증폭되는 현상이 존재하지 않기 때문에 해당 기지국별 측정치는 중계기의 영향 없이 기지국 신호만 수신되었다고 판단할 수 있게 된다. 따라서 기지국과 인접한 거리에 있는 높은 수신 신호를 갖는 측정치만을 이용하여 군집화를 수행하기 때문에 두 군집 중심 간 x축 거리는 약 50 m 이상으로 비교적 낮은 값으로 추정되는 것을 확인할 수 있다. 이를 통해 두 군집 중심에 대한 x축 간 거리 임계값은 약 100 m로 설정하여 중계기의 영향을 판단하는 기준으로 사용되었다.

매칭된 기지국에 대한 DB용 수집데이터에서 중계기의 전파 영향이 존재한다고 판단되면, 이상치 제거 알고리즘을 적용하도록 한다. 실시를 위해 Fig. 6과 같이 5곳의 측위 지점에서 획득한 데이터를 사용하였다. 그림에서 빨간색 점은 Fig. 4에서 수집된 측정치를 이용하여 신호정보가 생성된 기준점을 의미하고, 파란색 점은 미수집 영역에 해당하는 기준점을 나타낸 것이다. 각

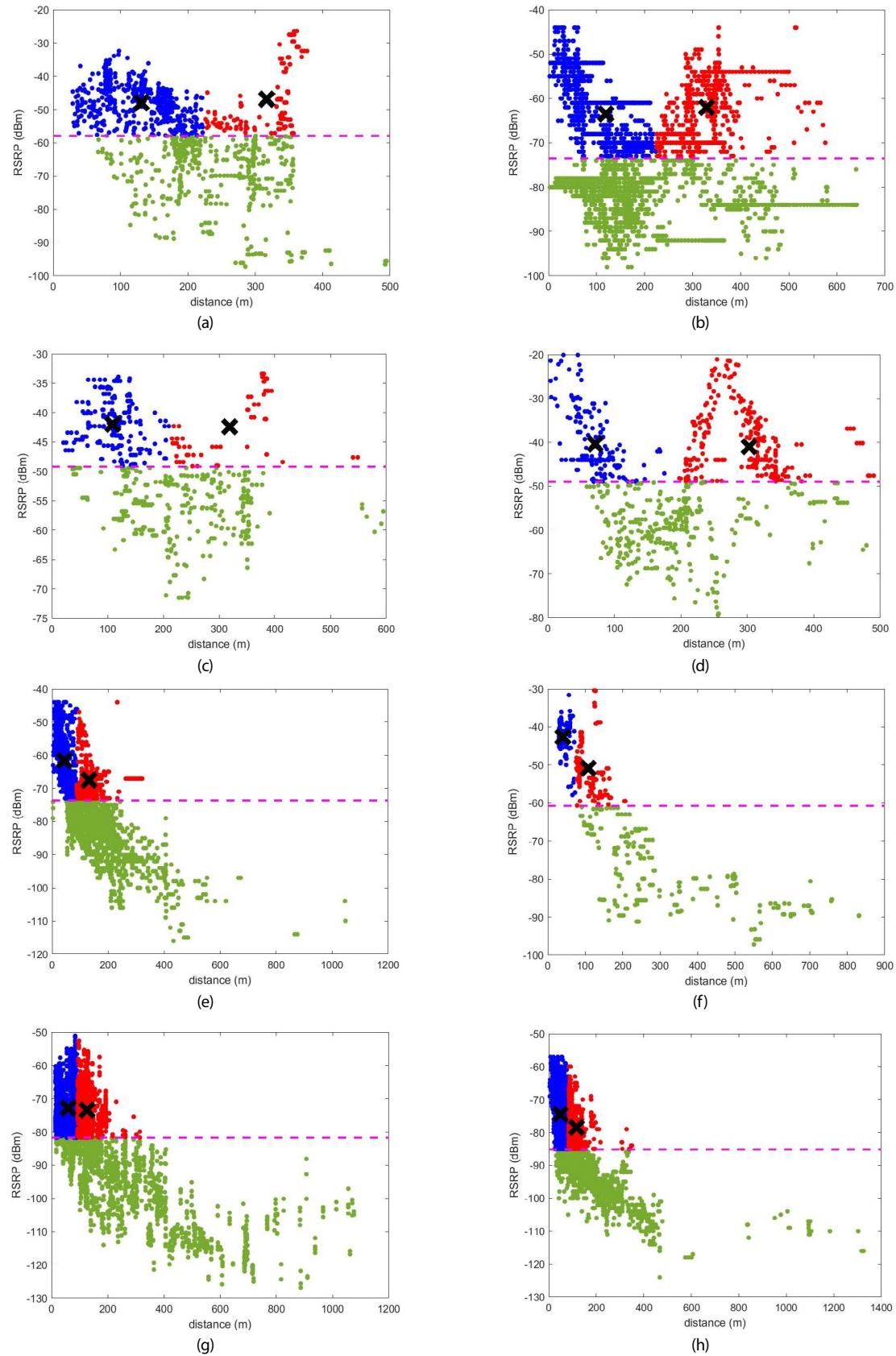


Fig. 5. Results of repeater signal detection. (a) PCI-174, (b) PCI-175, (c) PCI-211, (d) PCI-278, (e) PCI-12, (f) PCI-132, (g) PCI-144, (h) PCI-188

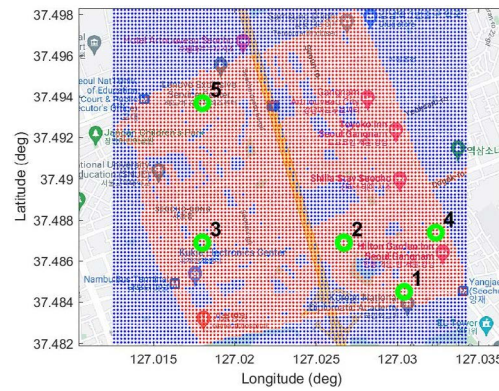


Fig. 6. Test points for positioning.

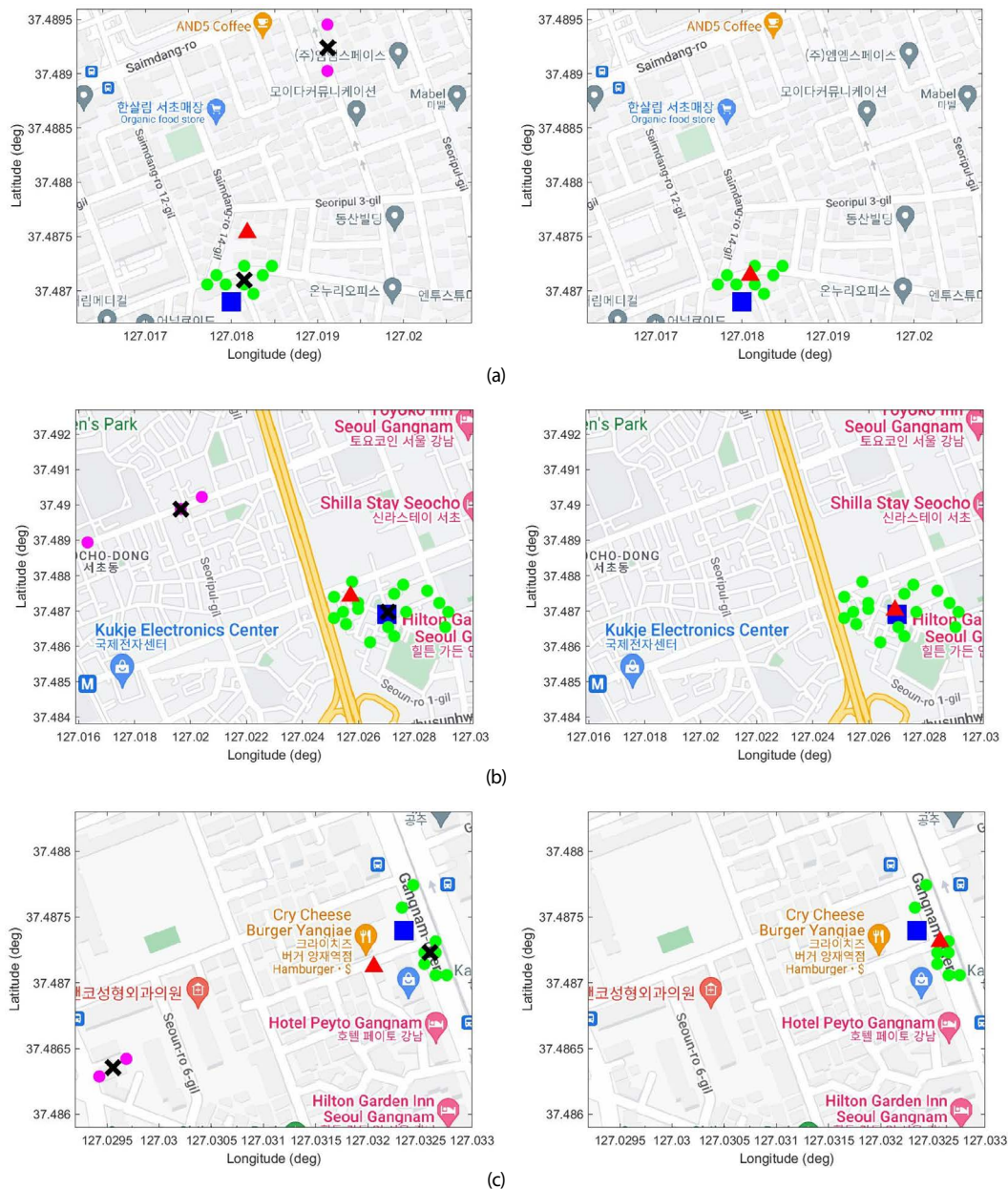


Fig. 7. Result of applying outlier removal technique. (left) before application, (right) after applying (a) Test 1, (b) Test 2, (c) Test 3, (d) Test 4, (e) Test 5

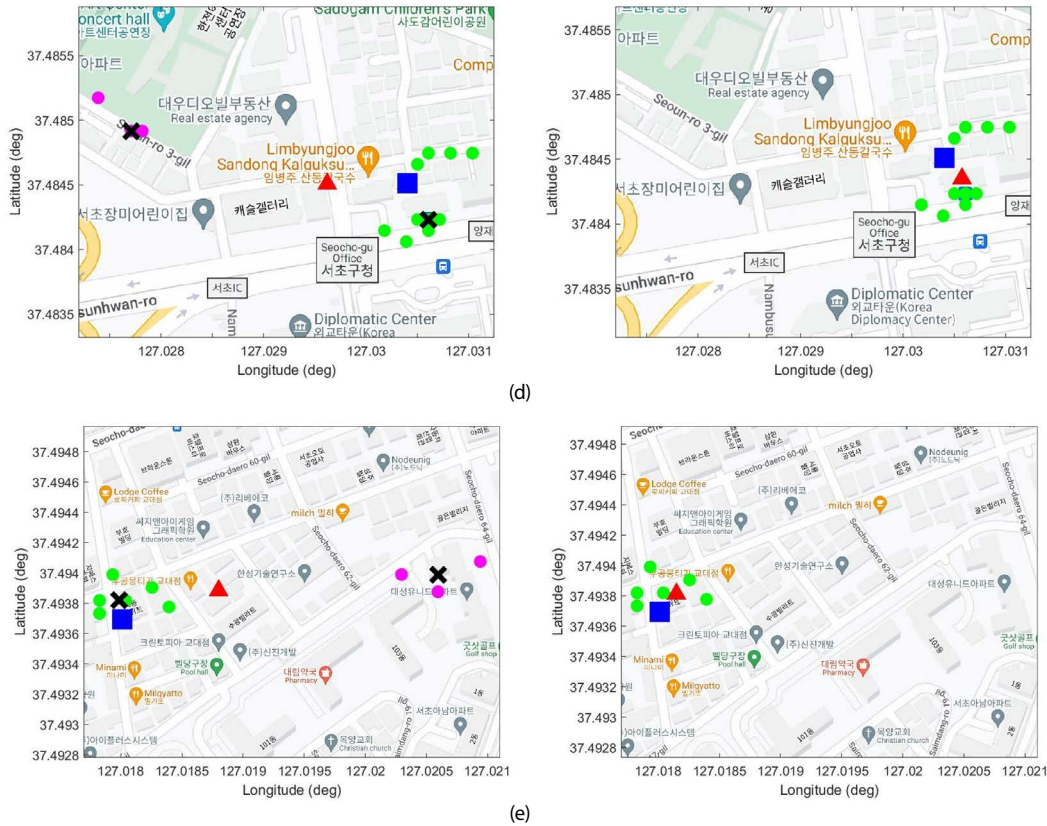


Fig. 7. Continued

측위 지점에서 90개의 데이터를 획득하였으며, 이상치 제거 알고리즘을 적용하기 전과 후의 측위 결과를 Fig. 7에서 각각 나타내었다. 이 때 해당 그림의 결과는 측위용 데이터에서 제안된 기법의 조건을 만족하는 일부 데이터를 이용하여 나타낸 결과이다. 먼저 왼쪽의 그림에서 파란색 네모는 측위 지점, 빨간색 세모는 추정 위치, 검정색 'x'는 군집 중심의 위치, 자주색 점과 초록색 점은 각각 2개의 군집으로 Clustering된 후보 기준점을 의미한다. 그리고 Table 2는 Fig. 7에서 나타낸 측위 결과를 위치 오차로 수치화 한 것이다. 먼저 후보 기준점을 보면 이상치로 추정되는 위치가 다수 포함된 것을 확인할 수 있다. 즉, 후보 기준점을 선정하는 과정에서 유사한 신호에 대해 중계기의 전파를 갖는 기준위치가 선택되는 것이다. 이로 인해 이상치에 해당하는 기준점도 일정 가중치를 적용하게 되므로, 최종적으로 출력된 위치정보의 추정 오차가 증가하는 것을 볼 수 있다. 선택된 후보 기준점은 대부분 참 위치 인근에 존재하며, 이상치의 영향으로 인해 군집 중심 간 거리는 커지게 된다. 따라서 기준위치의 분포도와 군집 중심 간 거리를 통해 이상치를 제거하도록 하였고, 이에 대한 결과를 오른쪽 그림에서 나타내었다. 중계기 신호 대한 기준점의 방향으로 가중되었던 추정된 위치정보는 해당 이상치를 제거함으로써 참 위치 인근으로 추정되는 것을 볼 수 있다. 실제로 Fig. 7에서 나타낸 5개의 측위 데이터를 기반으로 실험을 진행한 결과 위치오차가 평균 62.1 m 감소하는 것을 확인할 수 있었다. 또한 측위 지점마다 획득한 모든 데이터를 사용하여 제안된 기법을 적용하기 전과 후의 위치 오차를 Table 3에 나타내었으며, 평균 위치오차는

Table 2. Summary of positioning performance of the proposed technique.

Positioning area	Positioning error before and after applying the outlier removal algorithm	
	Before	After
1	73.3	29.0
2	151.9	16.2
3	53.6	25.9
4	75.1	26.0
5	71.6	17.8
Mean	85.1	23.0

약 26.6 m, 평균 표준편차는 약 33.1 m 감소함으로써 측위 성능이 향상됨을 확인할 수 있었다.

추후 중계기의 신호 전파 모델링을 통해 해당 신호를 정밀하게 검출할 수 있다면 기지국 기반 측위 성능이 더 많이 향상될 수 있을 것으로 기대된다.

5. CONCLUSIONS

본 연구에서는 LTE RSRP 정보를 이용한 WkNN 기반 Fingerprinting 측위 과정에서 이상치에 해당하는 기준위치를 제거하는 방법을 제안하였다. 측위 과정에서 매칭되는 기지국 정보가 적을 경우 신호의 유사도 크기에 따라 후보 기준위치가 선정된다. 그러나 차량 시험을 통해 도로상에서 획득된 DB용 수집 데이터는 기지국 신호만 측정되는 것이 아닌 중계기의 전파 영향

Table 3. Summary of positioning errors for all test data.

Positioning area	Positioning error before and after applying the outlier removal algorithm			
	Before		After	
	Mean (m)	Standard deviation (m)	Mean (m)	Standard deviation (m)
1	59.5	67.0	13.9	8.5
2	124.1	140.8	83.6	90.3
3	60.5	30.6	54.4	29.3
4	67.5	69.5	51.1	33.2
5	68.1	46.0	40.4	15.9
Mean	73.7	69.8	47.1	36.1

도 함께 포함될 수 있다. 이로 인해 중계기의 신호가 저장된 기준 점이 후보 위치로 다수 선택된다면 최종적으로 추정된 위치정보의 추정 오차는 크게 증가하게 된다. 따라서 본 논문에서는 매칭된 기지국의 DB용 수집데이터를 이용하여 중계기의 전파 유무를 판단하는 기법을 제안하였다. 먼저 기지국의 위치를 기준으로 측정지와 거리 각각 계산하고, 평균 신호보다 높은 데이터만을 이용하여 두 개의 군집으로 clustering을 수행한다. 그리고 군집 중심 간 x축 거리를 계산하고, 설정된 임계값을 기준으로 중계기의 영향을 판별한다. 이 후 중계기의 신호가 포함된다 판단되면 clustering 기법을 기반으로 이상치에 해당하는 기준점을 제거하도록 하였다. 제안된 기술의 성능은 서울시 서초 1,2동에서 획득된 데이터를 기반으로 검증하였다. 5곳의 테스트 지점에서 제안된 기법을 적용한 결과 평균 위치 오차가 약 26.6 m 감소되는 것을 볼 수 있었다. 또한 각 기지국별로 획득된 데이터를 이용하여 중계기의 전파 특성도 함께 확인할 수 있었다.

ACKNOWLEDGMENTS

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2019-0-01401, Multi-source based 3D emergency LOcalization using machine learning techniques).

AUTHOR CONTRIBUTIONS

Jae Uk Kwon contributed to the design and implementation of the algorithm and to the writing of the manuscript. Myeong Seok Chae contributed to provide information in writing the manuscript. Seong Yun Cho supervised the research and reviewed the manuscript as a project administrator. All authors discussed the proposed approach and results.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Alazab, M. & Gupta, M. 2022, Trust, Security and Privacy for Big Data (Boca Raton: CRC Press)
- Cho, S. Y. 2020, LTE signal propagation model-based fingerprint DB generation for positioning in emergency rescue situation, *Journal of Positioning, Navigation, and Timing*, 9, 157-167. <https://doi.org/10.11003/JPNT.2020.9.3.157>
- Cho, S. Y. 2022, Simultaneous LTE Signal Propagation Modelling and Base Station Positioning Based on Multiple Virtual Locations, *Sensors*, 22, 5917. <https://doi.org/10.3390/s22155917>
- Huang, H. S., Gartner, G., Krisp, J. M., Raubal, M., & Van de Weghe, N. 2018, Location Based Services: Ongoing Evolution and Research Agenda, *Journal of Location Based Services*, 12, 63-93. <https://doi.org/10.1080/17489725.2018.1508763>
- Ismail, A. H., Kitagawa, H., Tasaki, R., & Terashima, K. 2016, WiFi RSS Fingerprint Database Construction for Mobile Robot Indoor Positioning System, 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 9-12 Oct. 2016, Budapest, Hungary, pp.1561-1566. <https://doi.org/10.1109/SMC.2016.7844461>
- Khoo, H. W., Ng, Y. H., & Tan, C. K. 2022, Enhanced Radio Map Interpolation Methods Based on Dimensionality Reduction and Clustering, *Electronics*, 11, 2581. <https://doi.org/10.3390/electronics11162581>
- Kwon, J. U., Chae, M. S., & Cho, S. Y. 2022, CNN-based Adaptive K for Improving Positioning Accuracy in W-kNN-based LTE Fingerprint Positioning, *Journal of Positioning, Navigation, and Timing*, 11, 217-227. <https://doi.org/10.11003/JPNT.2022.11.3.217>
- Liu, S., de Lacerda, R., & Fiorina, J. 2022, Performance analysis of adaptive K for weighted K-nearest neighbor based indoor positioning, The 2022 IEEE 95th Vehicular Technology Conference: VTC2022-Spring will be held in Helsinki, Finland 19-22 June 2022. <https://doi.org/10.1109/VTC2022-Spring54318.2022.9860699>
- Xia, S., Liu, Y., Yuan, G., Zhu, M., & Wang, Z. 2017, Indoor Fingerprint Positioning Based on Wi-Fi: An Overview,

ISPRS International Journal of Geo-Information, 6, 135. <https://doi.org/10.3390/ijgi6050135>

Xu, S., Chen, C.-C., Wu, Y., Wang, X., & Wei, F. 2020, Adaptive Residual Weighted K-Nearest Neighbor Fingerprint Positioning Algorithm Based on Visible Light Communication, *Sensors*, 20, 4432. <https://doi.org/10.3390/s20164432>

Yuan, C. & Yang, H. 2019, Research on K-Value Selection Method of K-Means Clustering Algorithm, *Multidisciplinary Scientific Journal*, 2, 226-235. <https://doi.org/10.3390/j2020016>

navigation systems, filtering theory for linear/nonlinear systems, inertial motion detection, autonomous driving system, and location-based services.



Jae Uk Kwon received the B.S. degree from the Mechanical Automotive Engineering at Kyungil University and M.S. degree from the Department of IT Engineering at the same university in 2020, and 2022, respectively, where he is currently pursuing the Ph.D.

degree with the Department of IT Engineering. His current research topics include LTE-based fingerprinting for positioning, UWB-based wireless positioning, and navigation for autonomous vehicle.



Myeong Seok Chae received the B.S. degree from the Mechanical Automotive Engineering at Kyungil University and M.S. degree from the Department of IT Engineering at the same university in 2020, and 2022, respectively, where he is currently pursuing the Ph.D. degree with the Department of IT

Engineering. His current research topics include vehicular navigation system, pedestrian inertial navigation, and estimation theory.



Seong Yun Cho received the B.S., M.S., and Ph.D. degrees in Control and Instrumentation Engineering from Kwangwoon University in 1998, 2000, and 2004, respectively. From 2003 to 2004, he was an Assistant Researcher with Automatin and

System Research Institute, Seoul National University. He was a BK 21 Post-Doctoral Fellow with Seoul National University in 2004. From 2004 to 2013, he was with Electronics and Telecommunications Research Institute as a senior researcher. From 2008 to 2013, he was an Adjunct Professor with the University of Science and Technology. In 2013, he joined the faculty of the Mechanical Automotive Engineering at Kyungil University, where he is currently an associate professor. His current research topics include positioning and

