RadioLoc: Learning Vehicle Locations with FM Signal in All-Terrain Environments

Xi Chen[1], Qiao Xiang[2], Linghe Kong[3], Xue Liu[1]
[1]McGill University, [2]Yale University, [3]Shanghai JiaoTong University

***Abstract*—Vehicle localization service is a fundamental component of intelligent transportation systems. The widely used satellite navigation systems perform poorly in urban areas because the lines of sight to satellites are blocked by complex terrain characteristics, *e.g.*, buildings, elevated streets and interchanges. In this paper, we design RadioLoc, a novel system achieving accurate, efficient, all-terrain vehicle localization with two key design points. First, RadioLoc harvests the frequency modulation (FM) signal, which has a higher availability than satellite signal in complex terrains, as the signal source for localization. Second, RadioLoc integrates modern machine learning techniques into the processing of FM signals to efficiently learn the accurate vehicle localization in all-terrain environments. We validate the feasibility of FM-based vehicle localization and corresponding challenges and practical issues via field tests (*e.g.*, signal distortion, signal inconsistency and limited in-vehicle radio bandwidth), and develop a series of advanced techniques in RadioLoc to address them, including a new multipath delay spread filter, a reconstructive PCA denoiser, a tailored FM feature extractor, an adaptive batching technique and a frequency sweep technique. We implement a prototype of RadioLoc and perform extensive field experiments to evaluate its efficiency and efficacy. Results show that (1) RadioLoc achieves a real-time localization latency of less than 100 milliseconds; (2) RadioLoc achieves a worst-case localization accuracy of 99.6% even in an underground parking lot, and (3) the horizontal error of RadioLoc is only one sixth of a dedicated GPS device even when the vehicle is moving at a high-speed (*i.e.*, 80 km/h) in a complex highway scenario.**

I. Introduction

Vehicle localization is one of the most critical services in intelligent transportation systems (ITS), and the foundation of many ITS applications, such as navigation, electronic toll collection, traffic monitoring, emergency response and autonomous driving. Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS) [1], are the most widely used civilian vehicle localization systems. However, even being augmented by advanced technologies, the efficacy of GNSS is still constrained by an inherent limitation: GNSS require a clear line of sight from the vehicle to at least three satellites for accurate localization. The impact of this limitation is particularly magnified in urban environments, where the line of sights are blocked by a large amount of obstacles made of concrete and steels such as stack interchanges, multi-level garages, underground parking, street canyons, elevated roads and tunnels.

To cope with this limitation and achieve all-terrain vehicle localization, academia and industry have investigated the feasibility of many alternative signal sources. For example, Assisted GPS (A-GPS) [2], [3], [4] utilizes the assistance of cellular networks to provide localization service to smart phones under a partially blocked sky. However, the localization errors of A-GPS are substantially larger than those of dedicated GPS devices [4]. Systems using other signal sources (*e.g.*, WiFi [5], [6], [7], [8], acoustic signals[9] and visible light [10], [11]) provide a high accuracy for indoor localization. However, such signals are either less available or highly dynamic for vehicles in complex terrains.

In this paper, we design RadioLoc, a novel system that achieves accurate, efficient, all-terrain vehicle localization with two key design points. First, RadioLoc adopts the *FM radio signal*, a wireless signal highly available in all-terrain environments, as the signal source for localization. The FM signal is more advantageous over other signals (*i.e.*, satellite, cellular, WiFi, acoustic and visible light) for all-terrain vehicle localization, because it is free, highly available in complex urban terrains (*e.g.*, underground parking garages and tunnels), and requires no additional reception hardware on vehicles. Second, RadioLoc integrates modern machine learning techniques into the processing of FM signals, including adaptive sampling, profile feature extraction and location computation, to efficiently generate the accurate vehicle localization in all-terrain environments.

Even with all the advantages of FM signal, the previous studies on FM-based indoor localization [12], [13], [14], and the recent progress of modern machine learning theories and systems [15], [16], however, realizing accurate, efficient, all-terrain vehicle localization is still *non-trivial*. Through field experiments in Section II, we identify a series of unique challenges. First, the multipath richness of FM radio signal propagation generates large delay spreads, leading to significant *signal distortions* around some FM station frequencies. Second, even at the same location, the power offsets and interference levels of FM signals could vary significantly, due to the diversities of vehicle models, weather conditions, and the manual tuning of radio power gains by users, which leads to the *inconsistency of FM signal fingerprints*. Third, the *high mobility of vehicles*, as well as the *limited bandwidth of in-vehicle radios*, further degrades the localization accuracy. Fourth, the location of vehicles needs to be computed in real-time even when vehicles are moving at a high speed.

To address these challenges, in RadioLoc, we develop multiple advanced techniques (Section III and Section IV). In particular, RadioLoc detects and eliminates the signal distortions caused by delay spreads directly (instead of estimating the delay spreads themselves), with a Mahalanobis distance based filter. A reconstructive Principle Component Analysis (rPCA) denoising technique is then designed to further reduce the residual noises. Second, to cope with the inconsistency of FM signal fingerprints, RadioLoc extracts essential signal features that are immune to diverse vehicle models and changing weather. The variations in power offsets and interference levels are neutralized. Third, RadioLoc embraces an adaptive batching technique, which adjusts the data collection periods

Xi Chen and Qiao Xiang are co-primary authors.

2155-6814/19/$31.00 ©2019 IEEE
DOI 10.1109/MASS.2019.00058

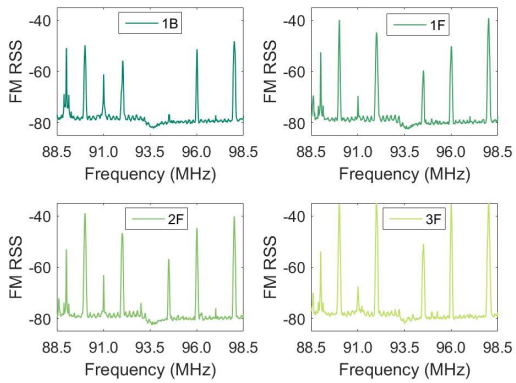


Fig. 1. FM RSS profiles of four positions.

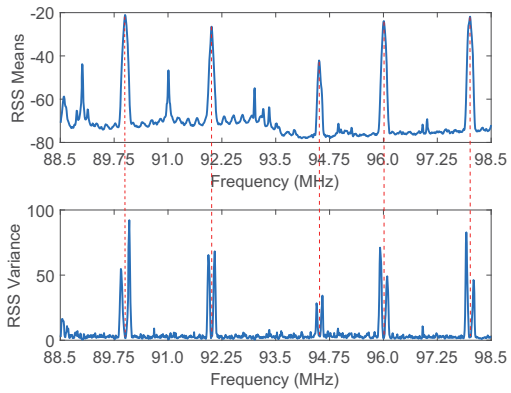


Fig. 2. Means and variances of RSS values at 1F over 200 samples.

according to the current vehicle velocity, and utilizes a frequency sweep technique to increase the system bandwidth for low-end radios. Fourth, RadioLoc adopts the random forest learning method to develop an algorithm that swiftly learns the accurate location of vehicles.

We implement a prototype of RadioLoc and perform extensive field experiments to evaluate its performance (Section V). Specifically, we conducted field experiments in two complementary scenarios with different terrains - a multi-floor parking building and a street section in an open neighborhood. In the first scenario, results show that RadioLoc achieves a worst-case localization accuracy of 99.6% in totally 18 locations scattering on four different floors (including one underground floor). A high-speed (*i.e.*, 80 km/h) test in the second scenario shows that RadioLoc lowers the horizontal errors to 16.7% of those given by a dedicated GPS device. And in both scenarios, RadioLoc achieves a localization latency of less than 100 milliseconds.

The **main contributions** of this paper are as follows.

- We design RadoiLoc, a novel FM-based vehicle localization system, which to the best of our knowledge is the first working system that achieves efficient, accurate, all-terrain vehicle localization;
- We identify the design challenges and practical issues of FM-based all-terrain vehicle localization through field tests, and develop a series of novel techniques to systematically address these issues;
- We fully implement RadioLoc and perform extensive field experiments to demonstrate the efficiency and efficacy of RadioLoc, in terms of localization latency and accuracy.

II. FM-based Vehicle Localization: Feasibility and Challenges

We identify the feasibility and corresponding challenges of FM-based all-terrain vehicle localization through field experiments in a four-floor parking building. The building has one basement $1B$ floor and three floors $1F$, $2F$ and $3F$ on the ground. We used USRP B210 boards as onboard FM radios, and left the built-in FM radios untouched to minimize the inconveniences to the volunteers.

A. Feasibility Experiments

We first parked an equipped vehicle on four different floors of the building. On each floor, this vehicle was parked at the same horizontal position. Hence, these four locations shared the same longitude and latitude, but had different altitudes. RSS profiles of FM signals were recorded at these locations. Here, an RSS profile is defined as a set of RSS values at different frequency points over a certain bandwidth. For each location, multiple samples of RSS profiles were recorded, and an average profile was calculated based on these samples.

The average RSS profiles of four different positions are illustrated in Fig. 1. In the RSS profiles, there are multiple peaks, each of which corresponds to a local radio station. It is shown that the peaks have different values and orders at different floor. These profiles are distinguishable from each other by analyzing their RSS peaks at the radio station frequencies. As such, we conclude that FM-based all-terrain vehicle localization is feasible.

B. Experimental Investigation of Challenges

Although we find that FM-based all-terrain vehicle localization is feasible, and some FM-based indoor localization systems are also recently developed, (*e.g.*, [12], [13], [14]). Many issues remain open when designing an FM-based vehicle localization system. To identify these issues, we further conduct several sets of field tests.

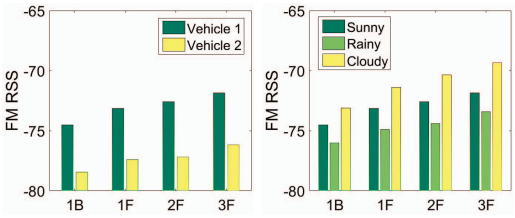
1) Impact of Multipath Delay Spread: The complicated urban terrain introduces a rich set of multipaths to the FM radio broadcasting, leading to a large delay spread in the FM signals. This spread in the time domain introduces random and unpredictable signal dispersion in the frequency domain. To analyze the impact of the delay spread, we further study the statistical details of FM RSS profiles. Fig. 2 takes $1F$ as an example, and depicts the means and variances of RSS values over 200 samples. An interesting observation is that, while the peaks of means appear at the frequencies of radio stations (we call them station frequencies in the rest of this paper), the peaks of variances are a bit off these station frequencies. This suggests that, while the centers of FM signals always locate at station frequencies, the shapes of them are varying dynamically. Since the vehicle was parked during recording, these variations should be mainly introduced by the multipath reflections of FM signals. The large and time-changing delay spreads were reflected as signal distortions in the frequency domain, leading to the large variances around radio station frequencies. Therefore, we have a unique challenge summarized as follows.

Challenge 1: The multipath richness of FM signals introduces severe delay spreads, which lead to previously undiscovered noises around station frequencies.

2) Impact of Vehicle Diversity: There are numerous ground vehicle models. Each of them may have a distinct placement

TABLE I
RELATIVE INTERFERENCE UNDER DIFFERENT WEATHER CONDITIONS

	Sunny	Rainy	Cloudy
Interference level	-78.2	-79.8	-76.3



(a) With different vehicles. (b) Under different weather.

Fig. 3. FM RSS variations due to different vehicles and weather conditions.

of radio antenna, a unique radio device, and different materials resulting in special signal propagation factors. Even vehicles from the same model may behave differently. To investigate the impact of vehicle diversity, we next employ a second vehicle from a different brand, and repeat the field tests at the same day as before. Fig. 3(a) compares the average RSS values in four different positions between two vehicles. Although both vehicles are parked at the same positions, their average RSS values are different. This suggests that the raw RSS profiles are not consistent across different vehicle setups. A closer investigation at the values reveals that there is a nearly constant power offset of $4.3dB$ between the results of two vehicles. Therefore, we summary the following challenge.

Challenge 2: The diversity in vehicle models may introduce variations in power-related features such as path losses and receiver-side gains. An almost constant yet unknown power offset is brought to FM signals, making their RSS profiles inconsistent.

3) Impact of Weather Conditions: Besides internal factors, the consistency of FM signals may also be affected by external factors. FM signal broadcasting could be sensitive to weather conditions. An example is the well-known rain fade phenomenon. However, we still need to understand the impact of weather conditions on FM RSS profiles and FM-assisted localization. Therefore, we repeat the tests with the first vehicle on different days under different weather conditions. Fig. 3(b) compares the average RSS values across three weather conditions. It is illustrated that the average RSS values vary with weather conditions. A further investigation reveals that the changes in interference level (Table I) contribute to these undesirable variations, leading to the following challenge.

Challenge 3: The diversity in weather conditions may introduce variations to the interference level, resulting in inconsistent FM profiles.

4) Other Practical Issues: In addition to the above challenges, there also exist other practical issues that need to be addressed.

Practical Issue 1: To cancel the white noise, it is required to collect multiple samples of RSS profiles at one location. However, when the vehicles are moving, it is impossible to gather multiple samples at exactly the same location. Instead, we may resolve to the collection of multiple samples within a small range. The size of this range may vary significantly with vehicle speed, which prevents us from obtaining consistent RSS profiles and FM fingerprints.

Practical Issue 2: The most important frequency components of RSS profiles are located at the frequencies of radio

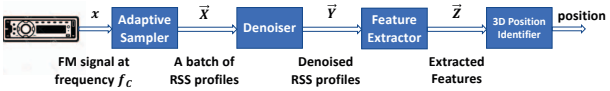


Fig. 4. Overview of RadioLoc.

stations, which are separated by hundreds of KHz. Low-end in-vehicle radios may not have a large enough bandwidth to cover an abundant number of radio stations in RSS profiles. As a result, the resolution and accuracy of FM localization may degrade.

Practical Issue 3: Vehicle localization has strong real-time and accuracy requirements. The accurate location of vehicles need to be computed in real-time even when vehicles are moving at a high speed.

III. RADIOLOC OVERVIEW

In this section, we give an overview of the architecture and workflow of RadioLoc, as illustrated in Fig. 4. RadioLoc utilizes its radio to gather the ambient FM signals x centering at the frequency f_C. This center frequency is determined and adjusted by the frequency selector. Suppose the bandwidth of these frequency signals x is B. Continuous signals are converted to digital signals, and are passed to the succeeding adaptive sampler.

Upon receiving each signal sample, the adaptive sampler calculates the RSS value on each frequency point, and obtains a RSS profile $\vec{x}$ covering the bandwidth B. These profiles are output with a rate of R samples per second. This rate is the output rate of digital samples with bandwidth B, and is different from the sampling rate of FM radio. We denote each sample of RSS profiles $\vec{x}$ as

$$\vec{x} = \{x_{f_1}, \ldots, x_{f_l} ..., x_{f_L}\}^T, \tag{1}$$

where x_{f_l} denotes the RSS value at the lth frequency point f_l, and L is the total number of frequency points in the frequency domain. The interval between two frequency points is thus B/L Hz. Every T seconds, the sampler creates an RSS profile batch with all the RSS profiles obtained during this period. Each batch is denoted as $\vec{X} = \{\vec{x}^{(1)}, \ldots, \vec{x}^{(i)}, \ldots, \vec{x}^{(N)}\}$, where $N = RT$, and the superscript i of $\vec{x}^{(i)}$ refers to the sequential number of profiles in one batch. Each batch is then output to the succeeding block. Note that the use of this batch technique is necessary to compensate the white noise across consecutive RSS profiles.

The denoiser is designed to eliminate the noises in each batch of RSS profiles $\vec{X}$, and outputs the denoised result as $\vec{Y}$. The noises to be cancelled include not only the white noise but also the signal distortions introduced by multipath delay spread. To avoid computationally heavy estimation of delay spread, the proposed denoiser employs a Mahalanobis distance based filter to detect and remove the distorted RSS values directly. It then applies a reconstructive PCA denoising technique to further reduce the residual noises.

The feature extractor aims to eliminate the FM signal inconsistency brought by the diversities of vehicles, radios, and weather conditions. It achieves this goal by extracting FM signal features that are irrelevant to these diversities. The extractor first normalizes each denoised profile $\vec{y}^{(i)}$ in the denoised batch $\vec{Y}$, so as to compensate the unknown power offsets brought by vehicle/radio diversity. It then calculates features of each normalized RSS profile, including the heights of peaks, the order of peaks, and a binary differential sequence

Algorithm 1 Mahalanobis distance based filtering

Input: A batch of RSS profiles $\vec{X}$ consisting of N samples, and the empirical percentage ρ of outliers.
Output: A batch of filtered profiles $\vec{Y}_M$.
Step 1: Calculate the distribution of variances in RSS values on each frequency point as follows.

1: **for** $l = 1$ to L **do**
2: $\quad \overline{x}_{f_l} = \sum_{i=1}^{N} x_{f_l}^{(i)}/N$.
3: $\quad v_l = \sum_{i=1}^{N} (x_{f_l}^{(i)} - \overline{x}_{f_l})^2/N$.
4: **end for**

Step 2: Calculate the Mahalanobis distances of all $\{v_l, l = 1, \ldots, L\}$ as follows.

5: $\mu = \sum_{l=1}^{L} v_l/L$.
6: $\delta = \left(\sum_{l=1}^{L}(v_l - \mu)^2/L\right)^{1/2}$.
7: **for** $l = 1$ to L **do**
8: $\quad$ Calculate the Mahalanobis distance as
9: $\quad \gamma_l = \left((v_l - \mu)(\delta^2)^{-1}(v_l - \mu)\right)^{1/2}$.
10: **end for**

Step 3: Remove outliers based on the Mahalanobis distances as follows.

11: Collect the largest $\lceil \rho L \rceil$ Mahalanobis distances, and put them into a set Γ.
12: $\gamma_{th} = \min\{\gamma_l | \gamma_l \in \Gamma\}$.
13: **for** $l = 1$ to L **do**
14: $\quad$ **if** $\gamma_l >= \gamma_{th}$ **then** $y_{f_l}^{(i)} = 0,\ \forall i = 1, \ldots, N$.
15: $\quad$ **else** $y_{f_l}^{(i)} = x_{f_l}^{(i)},\ \forall i = 1, \ldots, N$.
16: $\quad$ **end if**
17: **end for**

Step 4: Output $\vec{Y}_M = \{\vec{y}^{(1)}, \ldots, \vec{y}^{(N)}\}$.

Algorithm 2 rPCA denoising

Input: A batch of RSS profiles $\vec{Y}_M$ consisting of N samples.
Output: A batch of denoised RSS profiles $\vec{Y}$.
Step 1: Conduct standard PCA on $\vec{Y}_M$, and achieve the PC coefficient matrix, $\vec{C}$, the PC score vector $vecS$, and the estimated means $\vec{M}$ as follows.

1: The PC coefficient matrix: $\vec{C} = \{\vec{c}_1, \ldots, \vec{c}_L\}$, where $\vec{c}_l$ is of length L and is the coefficient vector of the lth PC, and $\{\vec{c}_l\}$ are in descending order of component variance.
2: The PC score matrix: $\vec{S} = \{\vec{s}_1, \ldots, \vec{s}_L\}$, where $\vec{s}_l$ is of length N and is the score vector the lth PC.
3: The estimated means: $\vec{m} = \{m_1, \ldots, m_L\}^T$, where m_l is the estimated RSS mean at the lth frequency point.

Step 2: Extend the vector of the estimated means into a matrix as $\vec{M} = \underbrace{\{\vec{m}, \ldots, \vec{m}\}}_{N}$.

Step 3: Filter the PC with the largest variance as follows.

4: The filter coefficient matrix: $\hat{C} = \{\vec{0}, \ldots, \vec{c}_L\}$, where $\vec{0}$ is a zero vector of length L.
5: The PC score matrix: $\hat{S} = \{\vec{0}, \ldots, \vec{s}_L\}$, where $\vec{0}$ is a zero vector of length N.

Step 4: Reconstruct the RSS profiles as $\vec{Y} = \hat{C}\hat{S}^T + \vec{M}$, and output $\vec{Y}$.

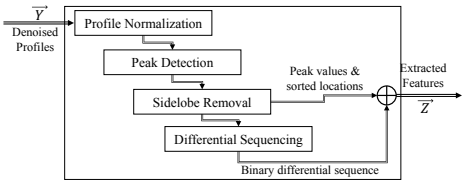


Fig. 5. Architecture of the feature extractor.

of the profile (to be defined in Section IV-B). These features are combined and weighted as a feature vector $\vec{z}^{(i)}$. The extractor processes all denoised profile in a batch, and outputs a batch of feature vector as $\vec{Z} = \{\vec{z}^{(1)}, \ldots, \vec{z}^{(N)}\}$. After extraction, the resulting feature vectors $\vec{Z}$ is sent to the 3D position identifier to compute the longitude, latitude and altitude of the vehicle.

IV. RadioLoc Design Details

After an overview of the basic workflow of RadioLoc, in this section we present the design details of RadioLoc and how these novel mechanisms address the challenges and practical issues identified in Section II.

A. The Denoiser

The denoiser consists of a Mahalanobis distance based filter and an rPCA denoising module.

1) Mahalanobis Distance based Filter: As previously illustrated in Fig. 2, an interesting observation of multipath delay spreads is that these spreads lead to large RSS variances around the station frequencies. This motivates us to pinpoint the distortions caused by delay spreads via detecting the frequency locations of high RSS variances. Concretely, we consider these high-variance points as outliers, and filter them out by detecting RSS values that are far away from the average distribution. We employ the Mahalanobis distance, as it quantifies the distance between a point and a distribution. The working procedure of the corresponding Mahalanobis distance based filter is described as Algorithm 1. Here, ρ is the empirical percentage of outliers, and is selected as 5% in this paper.

2) rPCA Denoising Module: To further clean the RSS profiles $\vec{Y}_M$, we develop a reconstructive PCA denoising technique. The corresponding rPCA denoising module analyzes the Principle Components (PCs) of the batch $\vec{Y}_M$, removes the most noisy one among these PCs, and reconstructs the profiles with the remaining ones as $\vec{Y}$. In this way, random noises and residual distortions are reduced. This procedure is described as Algorithm 2.

Note that this rPCA denoising technique is also able to remove the distortions brought by delay spread, since it can effectively eliminate high-variance elements from the profiles. However, it is hard to determine the PCs corresponding to the distortions in advance. Therefore, we develop the Mahalanobis distance based filter instead, and employ this rPCA denoising technique for other kinds of noises.

By applying the Mahalanobis distance based filter and the rPCA denoising technique, RadioLoc removes the distortions caused by delay spreads as well as other noises, and thus addresses **Challenge 1**.

B. The Feature Extractor

Fig. 5 presents the architecture of the feature extractor.

1) Profile Normalization: As investigated in Section II-B, the device diversity may introduce an unknown power offset to the FM signals, leading to inconsistency RSS profiles among different vehicles. In addition, drivers and passengers also may also adjust the power gains of in-vehicle radios, which also contributes to this unknown power offset. Such offsets are frequency non-selective. Therefore, the feature extractor minimizes the impact of these offsets by normalizing each RSS profile in $\vec{Y}$, and achieves the normalized profiles as $\overline{Z} = \{\overline{z}^{(1)}, \ldots, \overline{z}^{(N)}\}$, where $\overline{z}^{(i)}$ is the normalized version

Algorithm 3 Light-weight sidelobe removal

Input: A vector of detected peaks $\vec{P}$, their corresponding frequencies $\vec{F}$, a length of frequency window B_W, and a threshold ratio β.
Output: Sidelobe removal results $\hat{P}$ and $\hat{F}$.
Step 1: Take the first element in $\vec{P}$ as the current element, denoted with subscript c.
Step 2: Collect all the peaks whose frequencies are inside the range of $[f_c^p - B_W, f_c^p + B_W]$, and put them into set W.
Step 3: Check sidelobes in W.

```
for ∀z̃_l ∈ W do
    if z̃_l/z̃_c ≤ β then Remove z̃_l from P⃗, and remove f_1^p from F⃗
    end if
end for
```

Step 4: Repeat step 2 and step 3 for the next element in $\vec{P}$. Upon reaching the end of $\vec{P}$, collect the remaining elements in $\vec{P}$ and $\vec{F}$, conduct zero-padding at the ends of them to make sure their lengths equal K, and output the padding results as $\hat{P}$ and $\hat{F}$, respectively.

of $\vec{y}^{(i)}$ with a range of $[0, 1]$. This normalization is able to address Challenge 2 and Challenge 3 partially.

2) Peak Detection and Sidelobe Removal: In order to fully address these two challenges, we need to extract FM features that are irrelevant to the diversities of vehicles, radios, and weather conditions. To this end, we resolve to the appearances of different radio stations, as well as their relative order in RSS. The reasons are as follows. At a certain location, the number of radio stations and their frequencies inside a fixed bandwidth should be the same, regardless the reception hardware and weather conditions. In case that low-end radios lose a few stations to the noise, we further extract the relative order of these stations with respect to RSS values. In this way, high-end and low-end radios will always share most part in their orders of station RSS values.

To extract these features, the extractor first calculates the mean RSS value at each frequency point for all profiles in $\overline{Z}$:

$$\tilde{z}_{f_l} = \frac{1}{N}\sum_{i=1}^{N} y_{f_l}^{(i)}, l = 1, \ldots, L, \tag{2}$$

and denote $\tilde{Z} = \{\tilde{z}_{f_1}, \ldots, \tilde{z}_{f_1}, \ldots, \tilde{z}_{f_L}\}^T$. The extractor then finds the largest K peaks in $\tilde{Z}$ and their corresponding locations, by applying the peak detection algorithm proposed by Du *et al.* in [17]. The RSS values of the peaks are recorded in a descending order as $\vec{P} = \{\tilde{z}_1, \ldots, \tilde{z}_K\}^T$. The frequencies corresponding to these peaks are also recorded as $\vec{F} = \{f_1^p, \ldots, f_K^p\}^T$. If the number of detected peaks is less than K, the extractor performs zero-padding at the ends of both $\vec{P}$ and $\vec{F}$.

However, the peaks recorded in $\vec{P}$ do not always correspond to the main lobes of radio stations. They may include sidelobes that are introduced by the finite FFT in digital signal processing. Detect the sidelobes exactly is computationally complex. Therefore, we resolve to a light-weight algorithm to remove the sidelobes from $\vec{P}$ and $\vec{F}$, as described in Algorithm 3.

3) Binary Differential Sequencing: Besides the peaks, RadioLoc also leverages the shape of the profiles as another set of features. RadioLoc treats each profile as a series in frequency domain, and extracts its evolving trend along frequency. We utilize the differential sequence of each profile along frequency to capture how the RSS values evolve. The generation of this differential sequence is described as, for $l = 1, \ldots, L-1$,

$$\hat{d}_l = \begin{cases} 1, \tilde{z}_{f_{l+1}} - \tilde{z}_{f_1} \geq 0, \\ 0, \text{otherwise.} \end{cases} \tag{3}$$

We denote $\hat{D} = \{\hat{d}_1, \ldots, \hat{d}_{L-1}\}^T$. We adopt a binary differential sequence, because the magnitude information has already been captured in the peaks P. Using a binary format allows us to minimize the data storage overhead without much degradation on the localization performance.

4) Weighting the Features: After feature extracting, we obtain two kinds of features, the peak-relating features $\hat{P}$ and $\hat{F}$, and the trend-relating features $\hat{D}$. The number of peak-relating features are much less than the trend-relating features. However, peak-relating features may carry more important information about the RSS profiles. This information would be overwhelmed by the trend-relating features in the learning process, if we directly merge all features into one set. Therefore, we adopt a weighting method to make sure that all kinds of features will be treated equivalently regardless their numbers. The weight vector of weights is described as

$$\hat{W} = \{w_1^P, \ldots, w_K^P, w_1^F, \ldots, w_K^F, w_1^D, \ldots, w_{L-1}^D\}^T, \tag{4}$$

where $\{w_i^P\}$ are the weights for $\hat{P}$, $\{w_i^F\}$ are the weights for $\hat{F}$, and $\{w_i^D\}$ are the weights for $\hat{D}$. An example implementation of these weights is

$$w_i^P = 1/K, i = 1, \ldots, K, \tag{5}$$
$$w_i^F = 1/K, i = 1, \ldots, K, \tag{6}$$
$$w_i^D = 1/(L-1), i = 1, \ldots, l-1. \tag{7}$$

The final output of the whole feature extractor is

$$\vec{Z} = \{\hat{P}^T, \hat{F}^T, \hat{D}^T, \hat{W}^T\}^T. \tag{8}$$

All features in $\vec{Z}$ is uncorrelated to the diversities of vehicles, radios and weather conditions. Therefore, RadioLoc addresses **Challenge 2** and **Challenge 3**.

C. Practical Issues

We next present the novel mechanisms in RadioLoc to cope with the practical issues identified in Section II-B4.

Adaptive batching: The high mobility of vehicles is always a concern of localization and navigation systems. Fortunately, the impact of Doppler shift is small enough to be neglected in our design. A straightforward calculation tells us that the Doppler shift is 20Hz for a $216km/h$ vehicle speed at the $100MHz$ FM frequency band.

We consider another issue brought by the high speed of vehicles, and propose a practical solution for it. Consider the batching technique described in Section III, which packages the RSS profiles as one batch in every T seconds. Suppose there is a vehicle driving with the speed of v m/s. In this case, each batch of RSS profiles is recorded during a journey of vT meters. Depending on the variable v, the RSS profiles may correspond to very different physical distances, even if they were recorded from the same starting position. The RadioLoc system could suffer from these variations in the distance resolution. For example, it may compare real-time batches with a 20m distance to fingerprint batches with a 2m distance, and thus yields inconsistent results. To avoid this issue, we propose to adaptively adjust the period of each profile batch according to the vehicle's speed. Suppose the required resolution of distance is d (Without any further explanation, d is set to 3 meters in the rest of this paper.). Then the batch period is online adjusted as $T = d/v$ seconds. For instance, when the speed is 10 m/s (*i.e.*, 36 km/h), each batch will contain RSS profiles recorded in 0.3 seconds. When the speed increases to 20 m/s (*i.e.*, 72 km/h), the batch period reduces to 0.15

seconds. Both batches still cover the same distance of 3 meters. As such, *Practical Issue 1* is addressed.

Frequency sweep: The number of peaks (*i.e.*, radio stations) captured in each RSS profile determines on the bandwidth B of the received FM signals. For some low-end FM radios, the bandwidth B may be too small to cover an abundant number of radio stations. The localization accuracy could be unsatisfactory in this case. To overcome this issue, we propose to adopt frequency sweep to enlarge the bandwidth. Support the desired bandwidth is $B_d > B$, starting from the lowest frequency of f_C. Then the FM radio will switch its center frequency one by one to cover the following frequencies: $\{f_C$, $f_C + B$, $f_C + 2B$, ..., $f_C + \lceil B_d/B \rceil B\}$. Therefore, with a sweep of $\lceil B_d/B \rceil$ narrowband samples, we can construct a virtual wideband RSS profile by concatenating them on the frequency domain. Note that in this case the batch period is still the same as T seconds, but the number of wideband RSS profiles in each batch reduces to $RT/\lceil B_d/B \rceil$. As such, we address the *Practical Issue 2*.

3D position identifier: random forest based location learning Given the feature vectors $\vec{Z}$, the 3D position identifier uses them as the input to an identification function G. In RadioLoc, we adopt a random forest based learning algorithm for this function G, due to its accuracy and efficiency. We omit the details of the algorithm due to space limit. [1] The output $q = G(\vec{Z})$ is the index of the estimated position in a position dictionary D. Each position in the dictionary is stored as $(q, \vec{P}_q)$, where $\vec{P}_q$ is position q's 3D coordinate in the format of (longitude, latitude, altitude). The function G can take in $\vec{Z}$ with different number of samples (*i.e.*, the number of N can vary). Each vector $\vec{z}^{(i)}$ will output an estimated position q_i. Then the majority of $\{q_i\}$ is taken as the final identification output. In case of a draw, one of the majorities will be randomly selected as the output. As such, we address the *Practical Issue 3*.

V. Performance Evaluation

In this section, we evaluate RadioLoc with extensive field experiments in different terrain scenarios with different vehicle models, weather conditions and velocities.

A. Field Experiment Setup

1) Two Complementary Scenarios: The field experiments were conducted in two complementary real-life scenarios with different terrains.

Scenario 1 is a 3D scenarios used to evaluate the impact of diversities in vehicles and weather conditions. In order to achieve general conclusions, experiments in Scenario 1 may or may not have a clear view of sky, depending on the experiment positions. We choose a multi-level parking building to set up our experiments. This is because that, in the parking building, we can gather the ground truth data with 100% confidence. Moreover, we are able to repeat and reproduce the experiments exactly and safely without violating any traffic regulations.

The parking building has a total of four floors - three floors on the ground (denoted respectively as $1F$, $2F$ and $3F$) and one underground floor (denoted as $1B$). The height of each floor is 5 meters. On each floor, we select the same five horizontal positions ($P1$ to $P5$) as the reference points to gather FM data. We denote each reference point in Scenario 1 by its level and it horizontal position. For example, $P1$ on the second floor is denoted as $L2P1$. The horizontal floor plan of each floor is identical (except that $L3P2$ and $L3P4$ are both occupied by offices on the third floor), and is illustrated by Fig. 6. On each floor, $P2$ is at the entrance to the higher level, $P4$ is at the exit from the higher level, and $P3$ is between the entrance from the lower level and the exit to the lower level. $P2$, $P3$ and $P4$ are in a line, separated by 15 meters. $P1$ is 5 meters away from $P2$, and line $P1P2$ is perpendicular to line $P2P4$. The situation of $P5$ is similar to $P1$.

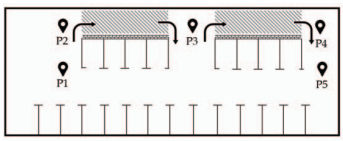


Fig. 6. Floor plan of parking building.

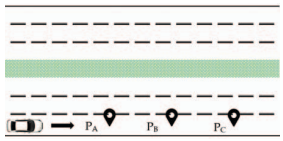


Fig. 7. Reference points in street.

Scenario 2 is used to evaluate the impact of vehicle speed. In order to concentrate on this issue and isolate others as much as possible, this scenario is set up in an open area with a clear view of sky. All field experiments in Scenario 2 are conducted at the same day. As illustrated by Fig. 7, we evaluate RadioLoc on a straight section of a 6-lane bi-directional street. There is no building (nor any sky-blocking object) within 100 meters of this section. In order to collect the ground truth for evaluation, we select three reference points (P_A, P_B, and P_C) locating on the midline of the rightmost lane. The distance between P_A and P_B, as well as that between P_B and P_C, is 5 meters. And the distance between P_A and P_C is 10 meters. During each run of experiments in Scenario 2, the same vehicle drives through this section with a constant speed, and passes the reference points in the order of P_A, P_B, and P_C.

Note that Scenario 1 and Scenario 2 are complementary to each other with different focuses. On one hand, Scenario 1 concentrates on the impact of environment diversities. The impact of speed is neutralized due to the nature that vehicles in a parking building are either static or driving slowly. On the other hand, Scenario 2 focuses on the impact of vehicle speed, and eliminates the diversities in vehicles and weather conditions by conducting experiments in the same day with the same vehicle.

2) Vehicles, Devices and System Setup: We employ two different vehicles in the experiments to evaluate the impact of vehicle diversity. The first vehicle (denoted as V1) was a Great Wall Haval M2 SUV. The second vehicle (denoted as V2) is a Nissan Rogue SUV. These two vehicles are owned and driven by two volunteering drivers, respectively.

In order to minimize the troubles that we bring to the volunteering drivers, we resolve to a portable and non-intrusive solution in collecting the FM signals. We adopt a USRP B210 board to create a software-defined FM radio, and connect it to a laptop for recording. This FM radio is mounted in front of the front passenger seat, next to the windscreen.

To collect the ground truth, we utilize the in-car Digital Video Recorders (DVRs) - the Jado D610s DVRs, on the rear-view mirrors of both vehicles. The ground truth locations at each time slot are extracted from the recorded videos. These videos are synchronized to the portable FM radio, so that we can compare the results of RadioLoc with the ground truth.

We also record GPS information with a GPS device - Garmin nuvi C265, which is also synchronized to the FM radio for comparison.

[1] The design of RadioLoc is modular such that different learning algorithms can be plugged in as G.

Day	Scenario	Weather	Vehicle	Time
1	1	sunny	V1, V2	9:00
2	1	sunny	V1	12:30
3, 4	1	cloudy	V1	9:30, 22:00
5, 6	1	rainy	V1	16:30, 22:00
7	2	sunny	V2	23:00

Fig. 8. The summary of data collection.

	L_{-1}	L_1	L_2	L_3	recall
L_{-1}	30.2%	69.7%	0.1%	0.0%	30.2%
L_1	35.1%	64.2%	0.7%	0.0%	64.2%
L_2	20.7%	0.5%	78.8%	0.0%	78.8%
L_3	0.0%	31.7%	21.5%	46.8%	46.8%
precision	35.1%	38.7%	77.9%	100.0%	

Fig. 9. Confusion matrix of floor identification without profile normalization.

Recall \ Training	Sunny	Rainy	Cloudy
Sunny	99.5%	73.1%	88.5%
Rainy	74.1%	99.2%	84.6%
Cloudy	85.2%	77.8%	99.6%

Fig. 10. Average recall of RadioLoc without feature extraction.

Method \ Altitude error	mean	max
GPS	9.5 m	15.0 m
RadioLoc with a standard denoiser	2.9 m	15.0 m
RadioLoc with the proposed denoiser	0.3 m	5.0 m

Fig. 11. Geometric errors in Scenario 1.

	L_{-1}	L_1	L_2	L_3	recall
L_{-1}	99.5%	0.5%	0.0%	0.0%	99.5%
L_1	0.7%	99.3%	0.0%	0.0%	99.3%
L_2	0.0%	0.5%	99.5%	0.0%	99.5%
L_3	0.0%	0.0%	0.7%	99.3%	99.3%
precision	99.3%	99.0%	99.3%	100.0%	

Fig. 12. Confusion matrix of floor identification with profile normalization.

Recall \ Training	Sunny	Rainy	Cloudy
Sunny	100%	99.6%	99.7%
Rainy	99.7%	100%	99.6%
Cloudy	99.7%	99.7%	100%

Fig. 13. Average recall of RadioLoc with feature extraction.

3) Data Collection: The ground truth, the FM signal measurement and the GPS information are collected in seven days with either one or both vehicles. The weather of each day is either sunny, rainy or cloudy. For each weather condition, we repeat the measurement twice in two different days. Moreover, data collections are conducted either in the morning, at noon, in the afternoon or at night. The weather conditions, the number of vehicles, the scenarios of data collection and the collection time are summarized in Fig. 8. Note that the numerical order of days in Fig. 8 is different from their calendar order.

4) Metrics: In this paper, we analyze the performance of localization systems from two different aspects.

On one hand, we evaluate whether a vehicle is correctly identified to the closest reference point. For many vehicular applications, identifying a vehicle to a reference point is as critical as (or even more important than) retrieving the exact coordination. For example, the ETC service needs to know not only the coordinations of nearby vehicles but also whether the vehicles are at the ETC gates. Also, even if errors exist, the navigation can still perform properly when vehicles are accurately located to critical references such as intersections and forks of roads, and levels of overpasses. Accordingly, the following metrics are adopted.

The **precision** of position identification for reference point i is defined as:

$$precision_i = \frac{tp_i}{tp_i + fp_i}, \tag{9}$$

where tp_i and fp_i are the numbers of true positives and false positives of reference point i, respectively. The average precision over all cases are then defined as:

$$\overline{precision} = \frac{1}{N}\sum_{i=1}^{N} precision_i, \tag{10}$$

where N is the number of reference points.

The **recall** of position identification of reference point i is defined as:

$$recall_i = \frac{tp_i}{tp_i + fn_i}, \tag{11}$$

where fn_i is the number of false negatives of reference point i. The average recall is then defined as:

$$\overline{recall} = \frac{1}{N}\sum_{i=1}^{N} recall_i. \tag{12}$$

In this paper, we consider the recall to be equivalent to the localization **accuracy**.

On the other hand, we are also interested in errors in the geometric space. A geometric error is defined as the Euclidean distance between the estimated coordination to the ground truth coordination. The average geometric **localization error** is then defined as:

$$\bar{\epsilon} = \frac{1}{N}\sum_{i=1}^{N}\sum_{i=j}^{N} P_{i,j} d_{i,j}, \tag{13}$$

where $P_{i,j}$ denotes the probability that vehicles at point i are identified to reference point j, $\sum_j^N P_i j = 1$, and $d_{i,j}$ denotes the Euclidean distance between i and j.

In addition to the metrics presented above, one may wonder how fast can RadioLoc localize the vehicle. We note that in all of our experiments, we observe that our prototype is able to compute the location of the vehicle in less than 100 milliseconds. This demonstrates that RadioLoc is able to provide real-time vehicle localization. We omit the detailed results on localization latency due to the space limit.

B. The Impact of Multipath Delay Spread

We first evaluate whether the proposed denoiser is able to address the severe multipath reflections. We use the test data collected in Scenario 1, and focus on the geometric localization accuracy. Fig. 11 compares GPS, RadioLoc with a standard denoising technique [18], and RadioLoc with the proposed denoiser in terms of geometric errors. It is shown that RadioLoc can indeed improve the 3D localization accuracy over the existing GPS system. Furthermore, the standard denoising technique still suffers from the severe multipath delay spread, leading to a maximum error of 15 meters (*i.e.*, wrong by 3 floors). The proposed denoiser is able to minimize the impact of delay spread, and significantly reduces both the mean and maximum errors (*i.e.*, wrong by at most 1 floor). Therefore, we conclude that RadioLoc is able to address **Challenge 1**.

TABLE II
CROSS VALIDATION OF IDENTIFICATION W/O PROFILE NORMALIZATION.

Data being used	V1 only	V2 only	V1 and V2
Worst-case precision	99.8%	99.9%	35.1%
Worst-case recall	99.7%	100.0%	30.2%

C. The Impact of Diversity in Vehicle Models

In this section, we analyze the impact of vehicle diversity, and evaluate whether RadioLoc can work consistently across different vehicle models. We evaluate the performance of altitude localization, and check whether the vehicles can be accurately located at the floor level. The data used for evaluation are collected in Scenario 1 during Day 1 with two vehicles. The measurement from both vehicles are mixed together. Recall that the feature extractor is proposed to address the vehicle diversity. Therefore, we compare RadioLoc systems with and without the profile normalization module. The confusion matricides are presented in Fig. 9 and Fig. 12, respectively.

It is shown in Fig.9 that, without RSS profile normalization, there will be a very high chance that vehicles are located to a wrong floor. The worst-case precision and recall are 35.1% and 30.2%, respectively. Considering the fact that there are totally four floors, these results are only slightly better than random guesses.

To investigate the cause of this, we further evaluate the worst-case precision and recall using the data from only one vehicle in Table II. We can see that the cross validation results are nearly perfect when the data of only one vehicle is adopted. However, the performance of floor identification degrades significantly when it is applied on the mixed data from both vehicles. This again confirms the existence of Challenge 2. The diversity of vehicles may bring variations to path losses and receiver gains, making the raw FM RSS inconsistent.

On the other hand, when the normalization module is applied, the diversity of vehicles can be appropriately addressed. As shown in Fig. 12, the performance of floor identification improves largely even with a mixture of data from both vehicles. The worst-case precision and recall are raised to 99.0% and 99.3%, respectively. Moreover, it is shown that the error in floor identification is at most one floor. Hence, we conclude that the profile normalization addresses the power offsets brought by different vehicle models, and thus solves **Challenge 2**.

D. Impact of Weather Conditions

In this section, we evaluate the performance of RadioLoc under different weather conditions in Scenario 1. The experiments are conducted in six different days with three different weather conditions, *i.e.*, sunny, rainy and cloudy. For each weather condition, experiments are repeated twice in two different days. For Day 1, data sets from both vehicles are employed.

We first try to identify the positions of vehicles to the 18 reference points on four floors. Again, a weighted random forest is adopted as an example identifier. The evaluation is conducted as follows. 1) We first conduct the training with the data of a weather condition that is different from those of testing data. Each run of the evaluation mixed all data of one weather condition as the training data set, and considers the rest as testing data sets. The classifier is trained with the training data set, and is then test separately on testing data sets. We denote the case of training on the sunny data set and testing on rainy data set as the Sunny-Rainy pair (S-R for short). Similarly, the rest pairs are S-C, R-S, R-C, C-S, and C-R, respectively. 2) For pairs where the weather conditions are the same (*i.e.*, S-S, R-R, and C-C), we will take the data from one day as the training data, and consider the data from the other day with same weather as the testing data.

After iterating through all training-testing pairs, we summarize the average recalls over all 18 reference points in Fig. 10. Note that the results in Fig. 10 are achieved using RSS profiles $\vec{Y}$ directly (*i.e.*, RadioLoc w/o feature extraction). In this case, when training and testing with the data from the same weather condition, the FM position identification can still perform quite well, even when the data are collected from two different days. However, when training on one weather condition and testing on others, the performance of position identification is not satisfactory. In the worst case, the vehicles have been located to wrong reference positions in 26.9% of the time. This suggests that the weather condition indeed brings variations to the FM fingerprints. Considering the fact that the profile normalization has been already applied, we can conclude that the fingerprint distortions happen not only in the scale but also in the shape of the signals.

To address this issue, we adopt the extracted features $\vec{Z}$ instead of the RSS profiles $\vec{Y}$. Field test results confirm that the extracted features are able to improve the accuracy of position identification. As illustrated in Fig. 13, the worst-case recall increases to 99.6% for training and testing on different weather conditions, and reaches 100.0% for training and testing on the same weather.

We further visualize the improvements brought by the feature extractor in Fig. 14. It is shown that the average geometric localization errors of GPS are always greater than 15 meters. These errors are reduced to less than 5 meters by RadioLoc w/o feature extraction. Moreover, using the extracted features helps RadioLoc lower the average errors to less than half a meter. Therefore, we conclude that the RadioLoc addresses **Challenge 3**.

E. Impact of Vehicle Speed

As discussed in Section II, the collected batches of FM profiles at the same reference point could be inconsistent under different speeds, undermining the performance of RadioLoc. To evaluate the impact of vehicle speed and the adaptive batching technique we proposed in Section IV-C to address this issue, we conduct multiple runs of tests in Scenario 2 with vehicle V2. At each run, V2 drives through three reference points, *i.e.*, P_A, P_B and P_C, in a straight line with a constant speed. Different speeds are adopted in different runs, ranging from $10km/h$ to $80km/h$. We first take the measurement data of $10km/h$ as the training data of a weighted random forest, and use the rest data to test the trained random forest. We then repeat this procedure for other speeds, and obtain the horizontal localization errors as shown in Fig. 15.

From Fig. 15, we can see that for all methods the localization errors increase with speed. Compared to GPS, RadioLoc largely reduces the localization errors in the high speed domain. However, when the speed is less than $40km/h$, RadioLoc with fixed batching period (Fixed BP) results in larger errors than those of GPS. This is caused by the inconsistency in FM fingerprints at different speeds. This issue is addressed by the use of adaptive batching period (Adaptive

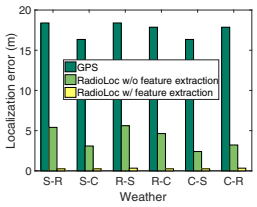


Fig. 14. Average 3D localization errors under different weather conditions in Scenario 1.

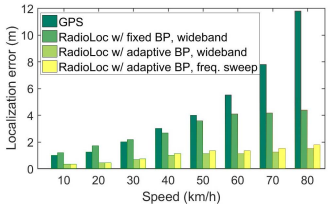


Fig. 15. Average horizontal localization errors under different velocities in Scenario 2.

BP). As demonstrated in Fig. 15, the use of adaptive batching period helps reduce the average localization errors by at least 57.7%. In the worst case, the average localization error is less than 2 meters. In the high speed scenario of 80 km/h, RadioLoc with adaptive batching period reduces the average error to 16.7% of the GPS error. Thus, we conclude that the proposed RadioLoc system addresses *Practical Issue 1*.

In addition, we further compare the use of wideband FM signals to the use of frequency sweep, so as to approach the limits of RadioLoc with low-end devices. The bandwidth of the wideband radio is 10 MHz, while the bandwidth of the narrowband radio is 500 KHz. It is illustrated in Fig. 15 that the performance degradation brought by frequency sweep is small. The worst-case average localization error is still less than 2 meters. Therefore, we conclude that the proposed RadioLoc system addresses *Practical Issue 2*, and can be safely applied to existing FM radios on vehicles.

VI. Related Work

Global positioning system. There has been a rich literature on how to improve the performance of GPS [2], [3], [4], [19], [20], including quality, accuracy, delay and etc. To support the functionality of GPS under poor satellite signal conditions, most GPS receivers nowadays are embedded with other radios, such as WiFi and cellular, to download the content of the GPS signal from assisted GPS (A-GPS) servers [2], [3], [4]. However, errors of A-GPS are significantly larger than those of standard GPS [4].

Localization using other signal sources. People also explored the potential and benefits of positioning with other wireless techniques, such as LTE, WiFi, radio frequency identification (RFID), acoustic signal and visible light [21], [6], [7], [22], [23], [24], [25], [26], [27], [28], [9], [8], [10], [9], [10], [11]. Common approaches adopted in these systems include the angle of arrival localization, the time of arrival localization and WiFi signal fingerprint and challenge state information localization. However, they mostly target on an indoor scenario and cannot support all-terrain 3D vehicle localization because (1) they usually require additional infrastructure and hardware modification and (2) they cannot cope with the highly dynamic vehicle environments.

Localization using FM signal. Recently, there has been a growing interest on FM-based localization methods [12], [13], [14]. Compared with other signal sources, FM radios consume less power and can cover a very large area. Though these systems demonstrate the benefits of FM-based localization. They are designed for indoor localization. FM-based all-terrain localization poses a series of unique challenges, such as the severe signal distortion caused by the rich multipaths in FM signal propagation, inconsistency of FM signals due to the diversities of vehicle models, radios, and weather conditions, the high mobility of vehicles, and the limited bandwidth of in-vehicle radios. To the best of our knowledge, RadioLoc is the first working system that systematically address these issues to achieve efficient, accurate, all-terrain vehicle localization.

VII. Conclusion

We design RadioLoc, a novel system that uses the highly-available FM signal as signal source and integrates modern machine learning techniques into the processing of FM signals to efficiently learn the accurate vehicle localization under all-terrain environments. A series of novel techniques are developed in RadioLoc to address the design challenges and practical issues in all-terrain vehicle localization. Field tests in real-life scenarios demonstrate that RadioLoc achieves a real-time localization latency with a worst-case accuracy of 99.6%.

References

[1] E.-R. Ahmed, "Introduction to gps: the global positioning system," in *Artech House communication series, Library of Congress, United States of America*, vol. 6, 2002.

[2] J. LaMance, J. DeSalas, and J. Jarvinen, "Assisted gps: a low-infrastructure approach," *GPS World*, vol. 13, no. 3, 2002.

[3] H. S. Ramos, T. Zhang, J. Liu, N. B. Priyantha, and A. Kansal, "Leap: a low energy assisted gps for trajectory-based services," in *UbiComp'11*.

[4] P. A. Zandbergen and S. J. Barbeau, "Positional accuracy of assisted gps data from high-sensitivity gps-enabled mobile phones," *Journal of Navigation*, vol. 64, no. 03, pp. 381–399, 2011.

[5] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "Csi-based indoor localization," *IEEE TPDS*, 2013.

[6] J. Xiong and K. Jamieson, "Arraytrack: a fine-grained indoor location system," in *NSDI 13*.

[7] S. Kumar, S. Gil, D. Katabi, and D. Rus, "Accurate indoor localization with zero start-up cost," in *MobiCom '14*.

[8] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wi-fi," *IEEE Transactions on Mobile Computing*, 2018.

[9] C. Zhang and X. Zhang, "Pulsar: Towards ubiquitous visible light localization," in *MobiCom'17*.

[10] T. Wei and X. Zhang, "Gyro in the air: tracking 3d orientation of batteryless internet-of-things," in *MobiCom'16*.

[11] L. Li, P. Xie, and J. Wang, "Rainbowlight: Towards low cost ambient light positioning with mobile phones," in *MobiCom'18*.

[12] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha, "Fm-based indoor localization," in *MobiSys '12*.

[13] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Mobicom '12*.

[14] S. Yoon, K. Lee, and I. Rhee, "Fm-based indoor localization via automatic fingerprint db construction and matching," in *MobiSys '13*.

[15] D. Kobak, W. Brendel, C. Constantinidis, C. E. Feierstein, A. Kepecs, Z. F. Mainen, X.-L. Qi, R. Romo, N. Uchida, and C. K. Machens, "Demixed principal component analysis of neural population data," *NIPS'16*.

[16] F. Nan, J. Wang, and V. Saligrama, "Pruning random forests for prediction on a budget," in *NIPS'16*.

[17] P. Du, W. A. Kibbe, and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching," *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, 2006.

[18] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons, 2006.

[19] D. Van Nee and A. Coenen, "New fast gps code-acquisition technique using fft," *Electronics Letters*, vol. 27, no. 2, pp. 158–160, 1991.

[20] H. Hassanieh, F. Adib, D. Katabi, and P. Indyk, "Faster gps via the sparse fourier transform," in *Mobicom '12*.

[21] Z. Yang, Z. Zhou, and Y. Liu, "From rssi to csi: Indoor localization via channel response," *ACM Computing Surveys*.

[22] H. Liu, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Accurate wifi based localization for smartphones using peer assistance," *TMC*, 2014.

[23] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices," in *MobiCom '14*.

[24] J. Wang and D. Katabi, "Dude, where's my card?: Rfid positioning that works with multipath and non-line of sight," in *SIGCOMM '13*.

[25] L. Li, G. Shen, C. Zhao, T. Moscibroda, J.-H. Lin, and F. Zhao, "Experiencing and handling the diversity in data density and environmental locality in an indoor positioning service," in *MobiCom '14*.

[26] S. Rallapalli, L. Qiu, Y. Zhang, and Y.-C. Chen, "Exploiting temporal stability and low-rank structure for localization in mobile networks," in *MobiCom'10*.

[27] D. Li, T. Bansal, Z. Lu, and P. Sinha, "Marvel: Multiple antenna based relative vehicle localizer," in *Mobicom '12*.

[28] X. Shen, Y. Chen, J. Zhang, L. Wang, G. Dai, and T. He, "Barfi: Barometer-aided wi-fi floor localization using crowdsourcing," in *MASS'15*.