UPS: Combatting Urban Vehicle Localization with Cellular-Aware Trajectories

Hua Xue*, Hongzi Zhu*, Siyuan Cao†, Shan Chang‡ and Jian Cao*
*Shanghai Jiao Tong University, China
Email: {howardsid, hongzi, cao-jian}@sjtu.edu.cn
†Purdue University, USA
Email: cao208@purdue.edu
‡Donghua University, China
Email: changshan@dhu.edu.cn

***Abstract*—Acquiring accurate location information of vehicles is of great importance. Global Positioning System (GPS) has been widely deployed and used to be the most convenient solution to outdoor localization. As more and more infrastructure such as elevated roads, tunnels and tall buildings is built, however, the ever-increasing complexity of urban environments makes vehicle localization especially in those urban canyons a new challenging problem. In this paper, we propose a novel scheme, called *UPS*, to tackle urban vehicle localization problem. Inspired by the observation from empirical study that the Received Signal Strength Indication (RSSI) values of cellular signals (e.g., GSM) perceived over a distance have ideal temporal-spatial characteristics for fingerprinting, UPS refines the location accuracy of a moving vehicle by matching its *cellular-aware trajectory*, which is an association between consecutive geographical positions and the corresponding wide-band GSM RSSI values, with a pre-constructed map. Moreover, UPS leverages large mobility of vehicles to construct large-scale maps. We implement a prototype system to validate the feasibility of the UPS design. We conduct extensive real-world experiments and results show that UPS can work stably in various urban settings and achieve an accuracy of 4.2 meters on average and 5.3 meters with a 90% precision.**

I. Introduction

Obtaining accurate location information of vehicles, especially in urban environments, is of great importance to many appealing applications. For instance, in the navigation application, the digital navigator can provide correct and timely instructions only when the accurate location information of a moving vehicle is acquired, particularly for those occasions when the road infrastructure is complex.

To solve the *urban localization problem*, which refers to obtaining the reliable and accurate location information of vehicles in urban settings, however, is very challenging due to the following three rigid requirements. First, an urban localization scheme should be *scalable* to the large area of a metropolis and the vast number of end users. Deploying new infrastructure or requiring users to have special devices is infeasible for the prohibitive cost. Second, such a scheme should achieve *timely* and *accurate* localization performance. For fast moving vehicles, large localization errors or delayed reports result in unpleasant user experience of LBS applications. Third, such a scheme should also be *resilient* to highly dynamic and complex urban environments such as different road types with varying traffic and different surrounding buildings and trees.

In the literature, existing outdoor localization techniques can be categorised into two classes, i.e., range-based and range-free. Range-based localization methods measure the distances or angles of a target object from given reference points and then perform trilateration or triangulation to obtain the estimated position of the target. ToA [1] can be used to determine the distance between the sender and the receiver. GPS is the most widely used ToA-based localization system. Many efforts try to make the tradeoff between the localization accuracy and power consumption on mobile devices [2] [3] [4] [5]. Differential Global Positioning System (DGPS) relies on a network of fixed, ground-based reference stations to improve the accuracy of GPS. Similarly, AoA [6] can be used to measure the angle between a pair of transceivers. Both of these two methods require specialized hardware and highly accurate synchronization, which make them less attractive for a large deployment. Pinpoint [7] improves the idea of ToA which needs no synchronization between devices and Time Difference of Arrival (TDoA) [8] also can be used to avoid clock synchronization. There are also efforts trying to model the relation between received signal strength (RSS) values and distances from senders. In practice, it is hard to build accurate and fine-grained propagation models in urban environment due to multipath, reflections and other effects.

Range-free localization does not rely on measurements of distances or angles. Cell ID has been used to get coarse-grained location information with an error of tens of meters [9]. Place Lab [9] is a example which estimates the location of the device referenced to the positions of radio beacons. CAPS [10] uses a cell ID sequence matching technique to estimate current position. There are also commercial systems such as Google's MyLocation [11] and Skyhook [12], requiring data of cell tower locations. Fingerprinting techniques are also introduced for localization. The key idea is to match location-dependent characteristic values to pre-collected data sets. RSS-based fingerprinting techniques are used both in outdoor [13] [14] [15] and indoor localization [16] [17] [18]. Channel State Information (CSI) is also used for indoor localization [19]. Note that although fingerprinting techniques can achieve better accuracy in indoor environments, it is hard to construct a reliable and fine-grained fingerprint map at a metropolitan scale.

978-1-5090-1328-9/16/$31.00 ©2016 IEEE

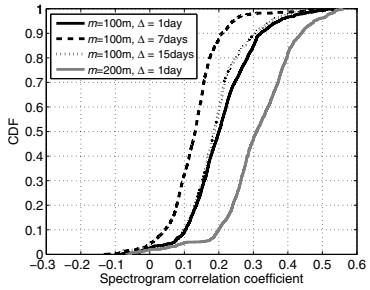


Fig. 1: Similarity of trajectories collected at different time.

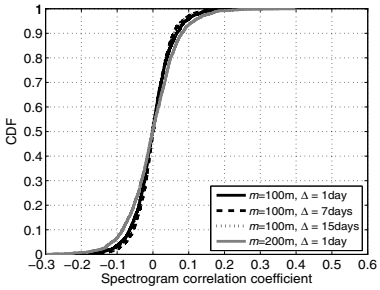


Fig. 2: Similarity of trajectories collected at different locations.

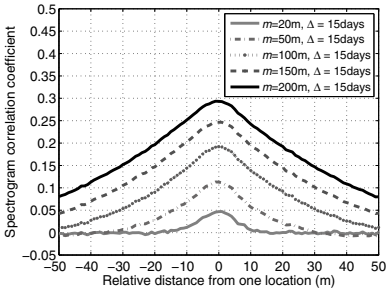


Fig. 3: Similarity of trajectories with a distance offset of k meters.

In this paper, to tackle the urban localization problem, we propose a novel scheme, called *Urban Positioning System* (UPS), which utilizes context information perceived along the trajectory of a vehicle. Based on the intensive analysis on a large-scale data set that we collected in Shanghai city, we have found that the *cellular-aware trajectory*, referring to the RSSI measurements of a wide band of GSM channels continuously measured along the movement of a high-speed vehicle, has very stunning temporal-spatial characteristics which makes it an ideal fingerprint for localization. Inspired by this insight, the core idea of UPS is to first use GPS or cell ID to determine a coarse location of a vehicle and then conduct a maximum-correlation search on a pre-constructed cellular-aware map with the cellular-aware trajectory of this vehicle to further refine the localization. We implement a prototype of UPS within a downtown area in Shanghai city of about six square kilometers, containing 54 road segments of about 30 kilometers, based on which we have conducted extensive field experiments. The experiment results demonstrate that UPS can achieve a high localization accuracy of 4.2 meters on average and 5.3 meters with a precision of 90%.

II. Empirical Study on Cellular-Aware Trajectories

In this section, we first describe the traces of cellular-aware trajectories that we collected in Shanghai city and then analyze the temporal-spatial characteristics of cellular-aware trajectories.

A. Collecting Cellular-Aware Trajectories

We choose GSM to study for three reasons. First, GSM has the ideal availability [14]. Second, GSM signals are very rich in urban settings. Third, the RSSI values of all channels can be measured.

Using OsmocomBB [20] and cheap GSM phones (i.e., Motorola C118), we can obtain a vector of RSSI values of 194 200kHz-wide R-GSM-900 channels, called a *power vector*, in very short time. We use 16 C118 phones mounted on the roof of an experiment vehicle to scan different channels in parallel. We conduct an extensive data collection campaign in a downtown area of six square kilometers in Shanghai city, where we experience various typical urban road conditions. With our experiment vehicle, the power vector of GSM band can be sampled at a rate of 10Hz. Meanwhile, the physical trajectory information of the vehicle (including the heading direction and distance information, see Subsection V-A for details) is also recorded. We traverse all 54 road segments within the selected area and collect the data twice a day for over one and a half months from May 4th to June 18th, 2014. The overall distance adds up to about 2,700km.

B. Temporal Stability of Cellular-Aware Trajectories

Considering the dynamic urban environment, such as trees and vehicle traffic, we first examine whether cellular-aware trajectories is temporally stable. In specific, we denote matrix $\mathbb{S}^{t,l,m} = (X_1^{t,l,m}; X_2^{t,l,m}; \cdots; X_n^{t,l,m})$ as a cellular-aware trajectory of n channels wide and m meters long, collected from location l at time t, where $X_i^{t,l,m} = (x_i^{t,l,1}, x_i^{t,l,2}, \cdots, x_i^{t,l,m})$, denoted as a *power strip*, are the RSSI values of channel i measured at time t from location l over a distance of m meters.

We measure the similarity between the power strips of channel i obtained on the same road segment but at different time t_1 and t_2 using Pearson's correlation coefficient as follows,

$$r_{X_i^{t_1,l,m} X_i^{t_2,l,m}} = \frac{\sum_{j=1}^{m}(x_i^{t_1,l,j} - \overline{X_i^{t_1,l,m}})(x_i^{t_2,l,j} - \overline{X_i^{t_2,l,m}})}{\sqrt{\sum_{j=1}^{m}(x_i^{t_1,l,j} - \overline{X_i^{t_1,l,m}})^2}\sqrt{\sum_{j=1}^{m}(x_i^{t_2,l,j} - \overline{X_i^{t_2,l,m}})^2}}, \tag{1}$$

where $\overline{X_i^{t,l,m}}$ is the average of $x_i^{t,l,j}$ for all $j \in [1,m]$. We then calculate the *trajectory correlation coefficient* (TCC) to measure the linear dependence between two cellular-aware trajectories collected on the same road but at different time t_1 and t_2 as follows,

$$r_{\mathbb{S}^{t_1,l,m}\mathbb{S}^{t_2,l,m}} = \frac{1}{n}\sum_{i=1}^{n} r_{X_i^{t_1,l,m} X_i^{t_2,l,m}}. \tag{2}$$

Note that cellular-aware trajectories obtained on the same road segment are geographically aligned before analysis (see Subsection IV-A for details). We randomly choose 500 cellular-aware trajectories of 100 and 200 meters long, respectively, from the trace collected during the first two weeks. For each trajectory, we calculate the TCC between this trajectory and trajectories collected at the same place but collected one day, one week and 15 days later, respectively.

Figure 1 plots the cumulative distribution functions (CDFs) of the TCC results. We have three observations. First, the cellular-aware trajectories have some degree of temporal stability. The average TCC value of 100m trajectories is above

0.1 in the figure. Second, this insignificant temporal stability will hardly fade away along time. For example, the TCC values calculated with a time difference of one week is smaller than that calculated with a time difference of 15 days. Last, increasing the trajectory length can help maintain the temporal stability. For instance, the average TCC value has increased to 0.3 when the trajectory length increases from 100m to 200m.

C. Geographical Uniqueness of Cellular-Aware Trajectories

The next step is to check whether cellular-aware trajectories are geographically unique, which refers to that trajectories collected from different locations within a sufficiently large area should be distinctive. To this end, we calculate the TCC between $\mathbb{S}^{t_1,l_1,m}$ and $\mathbb{S}^{t_2,l_2,m}$, which denote two trajectories of the same size but collected from different location l_1 at time t_1 and l_2 at time t_2, respectively.

We randomly choose 500 cellular-aware trajectories of 100 and 200 meters, respectively, from the trace collected during the first two weeks. For each trajectory, we randomly select 100 trajectories of the same length but at different locations from the trace collected one-day, one-week and 15-day later, respectively, and calculate the TCC. Figure 2 plots the CDFs of the results. We have two main observations. First, the cellular-aware trajectories have excellent geographical uniqueness. For example, about 95% of the TCC values are less than 0.1 over all time. Second, by comparing these CDFs with those presented in above subsection, an obvious gap can be seen and the gap can be enlarged by increasing the trajectory length. This major insight implies that cellular-aware trajectories can be used as fingerprints for localization.

D. Fingerprint Resolution Analysis

Besides the capability for fingerprinting, we analyze the resolution of cellular-aware trajectories, which refers to the minimum distance offset over which two trajectories can be distinguished. We randomly choose 500 trajectories of different length varying from 20m to 200m, respectively, from the trace. For each trajectory $\mathbb{S}^{t_1,l_1,m}$, we calculate the TCC using $\mathbb{S}^{t_1,l_1,m}$ and $\mathbb{S}^{t_2,l_2,m}$, where location l_2 is taken k meters away from location l_1 for $k \in [-50, 50]$.

Figure 3 shows the average TCC as a function of the offset distance k. It can be seen that, even for short trajectories, the TCC function is convex with the maxima appearing at the distance offset of zero, which implies that, theoretically, a target trajectory can be accurately localized by searching for the maximum TCC on a reference trajectory.

III. UPS System Overview

Despite of the insight that cellular-aware trajectories are excellent fingerprints for urban localization, there are still two main challenges in building a feasible urban localization system utilizing cellular-aware trajectories: 1) how to construct a calibrated and fine-grained fingerprint map, i.e., a map of reference cellular-aware trajectories, at a large scale; 2) how to deal with uncertain conditions of urban environments in accurately localizing individual vehicles. In UPS, we integrates two techniques that respectively tackle the above challenges: *fingerprint map construction* and *online vehicle localization*. The system architecture of UPS is illustrated in Figure 4.

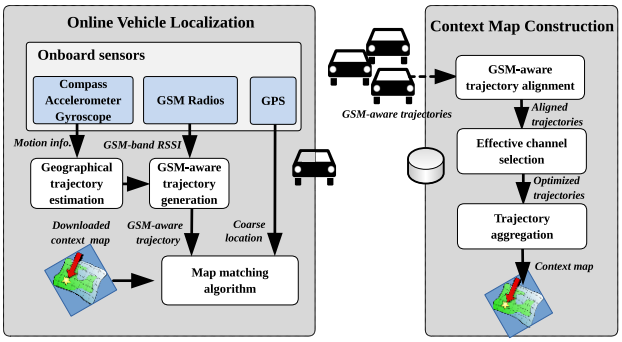


Fig. 4: System architecture of UPS.

Fingerprint map construction. The map construction is conducted at a data center. As one single measurement of a GSM-aware trajectory obtained on a road segment may contain noise, in the initial stage of UPS, we recruit dedicated vehicles to collect GSM-aware trajectories on each road segment for multiple times and aggregate those GSM-aware trajectories obtained on the same road to form the final fingerprint map in the following three steps. First, considering the diversity of each physical trajectories (e.g., driving on different lanes or changing lanes), we choose a "preferred" GSM-aware trajectory as the template and then align all other GSM-aware trajectories to it adopting the DTW technique. Second, the most effective GSM channels are selected for use based on their temporal stability and geographical uniqueness. Last, aggregation schemes can be applied to all aligned GSM-aware trajectories to finalize the initial fingerprint map for downloading and use.

Online vehicle localization. To estimate its location, a vehicle first utilizes its onboard motion sensors such as the accelerometer, gyroscope, and digital compass to estimate its physical trajectory, and continuously measures the power vector of the GSM band at the same time. By combining the measured power vectors to its physical trajectory according to time, it then generates its own GSM-aware trajectory. Finally, with a pre-downloaded fingerprint map and its GSM-aware trajectory, the vehicle conducts a map matching algorithm within a restricted area centred at the coarse location determined by GPS or Cell ID on the fingerprint map to determine its current location.

IV. Fingerprint Map Construction

A. Cellular-Aware Trajectory Alignment

In the initial stage, UPS deliberately collect a few cellular-aware trajectories on each road segment and align them before further processing.

Selecting the template cellular-aware trajectory. Given a set of cellular-aware trajectories obtained on a road segment $\{\mathbb{S}^{t_1,l,m_1}, \mathbb{S}^{t_2,l,m_2}, ..., \mathbb{S}^{t_k,l,m_k}\}$, we align the physical shapes of each pair of cellular-aware trajectories, adopting the DTW algorithm [21]. The algorithm will produce an optimized

match between two trajectories $\mathbb{S}^{t_i,l,m_i}$ and $\mathbb{S}^{t_j,l,m_j}$ with an minimized cost value $D(m_i, m_j)$ which is the summation of Euclidean distances between the matching locations. We select the cellular-aware trajectory $\mathbb{S}^{t_*,l,m^*}$ which has the minimum $\sum_{x=1}^{k} D(m^*, m_x)$ as the template trajectory since it has the least deviation of physical shape from all others.

Aligning cellular-aware trajectories to the template. Given the template $\mathbb{S}^{t_*,l,m^*}$, we align each cellular-aware trajectory $\mathbb{S}^{t_i,l,m_i}$, $i \in [1,k]$, with the template using DTW. As a result, we can get the correspondences of each location on $\mathscr{T}^{t_i,l,m_i}$ to one or many locations on $\mathscr{T}^{t_*,l,m^*}$. We generate a new cellular-aware trajectory $\widetilde{\mathbb{S}}^{t_i,l,m^*}$ by associating each location of the template with an averaged GSM power vector obtained from those measurements physically aligned to it.

B. Effective Channel Selection

Different GSM channels may contribute differently to the effectiveness of fingerprinting. We sort the channels and select the most effective ones.

In specific, for each pair of aligned cellular-aware trajectories $\widetilde{\mathbb{S}}^{t_1,l_k,m^*}$ and $\widetilde{\mathbb{S}}^{t_2,l_k,m^*}$ obtained on road segment l_k (the starting point of the road segment) at time t_1 and t_2, respectively, we can extract two power strips of channel i, denoted as $X_i^{t_1,l_k,m^*}$ and $X_i^{t_2,l_k,m^*}$, from the corresponding cellular-aware trajectories. The temporal stability of channel i on road segment l_k can be calculated with (1). We define the temporal stability utility of channel i, denoted as $u_i^{stability}$, as the average correlation coefficient over all road segments

$$u_i^{stability} = \frac{1}{K}\sum_{x=1}^{K} r_{X_i^{t_1,l_x,m^*} X_i^{t_2,l_x,m^*}} \tag{3}$$

where K is the total number of road segments. Similarly, we define the geographical uniqueness utility of channel i, denoted as $u_i^{uniqueness}$, as the average correlation coefficient between two power strips of channel i extracted from any pair of aligned cellular-aware trajectories obtained on different road segments

$$u_i^{uniqueness} = \frac{2}{K(K-1)}\sum_{j\neq k} r_{X_i^{t_1,l_j,m} X_i^{t_1,l_k,m}}, \tag{4}$$

where K is the total number of road segments. Finally, we define the utility of channel i as

$$u_i = u_i^{stability} - u_i^{uniqueness}. \tag{5}$$

We randomly choose 1,000 cellular-aware trajectories of 300 meters long from trace collected on May 4th and, for each channel (idle channels with index number ranging from 955 to 1023 are discarded), we calculate the utility of that channel with trace collected on May 18th. Figure 5 plots the temporal stability utility and geographical uniqueness utility as a function of channel index ranging from 0 to 124. It can be seen that the channel utility varies with different channels. In our prototype implementation, we select 41 channels with the utility higher than 0.2 (as illustrated by the dotted rectangles in the figure) as effective channels for urban localization. In practice, effective channel selection can be conducted for each area of interest at the data center based on the particular cellular-aware trajectories collected in that area.

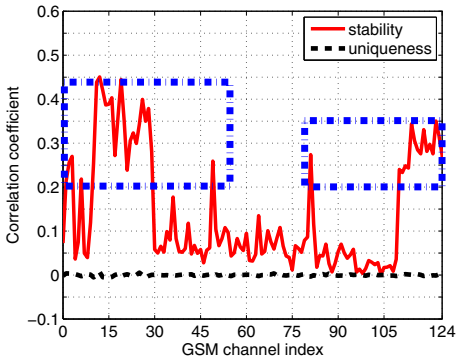


Fig. 5: Selecting effective channels based on channel utility.

C. Trajectory Aggregation

Given the set of aligned and channel-selected cellular-aware trajectories $\{\mathbb{S}^{t_1,l,m^*}, \mathbb{S}^{t_2,l,m^*}, ..., \mathbb{S}^{t_k,l,m^*}\}$, an aggregated cellular-aware trajectory is generated by averaging the GSM power vectors associated with each location on the aligned trajectories. More sophisticated aggregation schemes such as clustering and selective average (the outlier estimates are discarded before the rest estimates are averaged) can be adopted. After aggregating aligned cellular-aware trajectories on all road segments of interest, the fingerprint map is constructed.

V. Online Vehicle Localization

A. Physical Trajectory Estimation

For UPS, it is fundamental to obtain the accurate geographical trajectory information of a vehicle in terms of its distance that recently traversed.

Perceiving the heading direction. By aligning the coordinate system of the digital compass with that of the vehicle using the scheme proposed in work [22], the vehicle can utilize the onboard compass to get the strength and the direction of the magnetism of the earth by calculating the sum of magnetization readings along the x- and y-axis of the compass. As a result, the heading direction of the vehicle can be derived as the angle between the heading direction of the vehicle and the direction of the magnetism of the earth. In addition, we also use the onboard gyroscope to obtain the changes of heading direction.

Estimating the traversed distance. Given the instant velocity information of the vehicle, the traversed distance can be calculated as the integral of instant velocity values over time. Though it is straightforward to obtain the accurate velocity information via a dedicated OBD interface, inertial sensors like accelerometers and gyroscopes can also be used to estimate instant velocity [22].

With the heading and distance information, the vehicle estimates its m-meter geographical trajectory $\mathscr{T}^m$ as a vector of $m+1$ elements. Each element is a tuple (θ_i, t_i) for

$i \in [0, m]$, where θ_i and t_i represent the heading angle and the timestamp at the ith meter on the trajectory.

B. Cellular-Aware Trajectory Generation

A vehicle can deploy one or multiple GSM radios to measure the RSSI value of each GSM channel while moving. For each element (θ_i, t_i), $i \in [0, m]$, of a geographical trajectory $\mathscr{T}^m$ of the vehicle, the power vector $V^{t_i} = (x_1^{t_i}, x_2^{t_i}, \cdots, x_n^{t_i})$ measured over n channels during time interval of $[t_{i-1}, t_i]$ can be associated, forming the corresponding cellular-aware trajectory $\mathbb{S}^m$. It should be noted that, as it takes time to scan GSM channels, when the vehicle moves fast, it is possible that some channels (referred to as *missing channels*) within a power vector at a particular location are not measured. In this case, missing channels cause blanks with no valid RSSI values in the resolved cellular-aware trajectory.

C. Map Matching Algorithm

According to the fingerprint resolution analysis, a vehicle can be localized by searching for the maximum TCC between its cellular-aware trajectory and the reference cellular-aware trajectory on a fingerprint map.

In specific, the vehicle picks an m-meter cellular-aware trajectory of its own, $\mathbb{S}^m$, and searches on the map for one trajectory that has the maximum similarity in terms of TCC with $\mathbb{S}^m$, i.e.,

$$\arg\max_{l} r_{\mathbb{S}^m \mathbb{S}'^{l,m}}, l \in R^{l_0,d}, \tag{6}$$

where l is an arbitrary location within the disk region $R^{l_0,d}$ with the radius d, $\mathbb{S}'^{l,m}$ is a cellular-aware trajectory of m meters picked from the map starting from location l, and l_0 denotes the central location of the disk. The solution to (6), l^*, referred to as a *location reference point* (LRP), is treated as the best location estimation where $\mathbb{S}^m$ might start. In UPS, a sliding window of m meters is used to search for l^* road by road in $R^{l_0,d}$. With a LRP, the current location of the vehicle can be easily obtained by adding the distance travelled since that LRP.

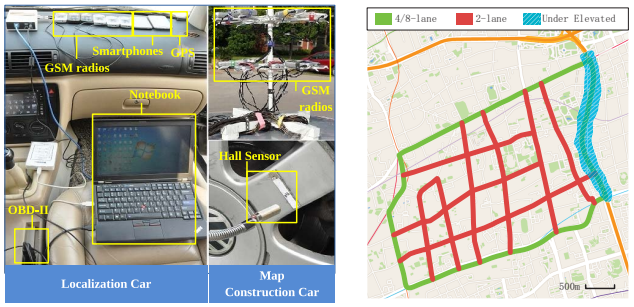
In practice, GPS location reports and a maximum GPS error can be used as the central location l_0 and the searching range d, respectively. Since vehicles are restricted to move on roads, given the limited area of $R^{l_0,d}$, the number of roads needed to be searched in that region is very limited.

VI. Performance Evaluation

A. Prototype Implementation

We implement a prototype system using two cars, i.e., one for map construction and the other for localization, as shown in Figure 6(a).

Online localization: We mount the different groups of Motorola C118 cellphones (working as GSM radios) placed at different locations on the localization car to scan the GSM band in parallel. In addition, we use two smartphones (i.e., one HTC S720t and one Samsung Galaxy S4) to estimate the heading direction of the car with their motion sensors, an OBD-II interface to read the instant speed of the car, and a high-performance GPS module.



(a) Experimental cars (b) Demonstration area

Fig. 6: The prototype implementation of UPS.

Map construction: Similarly, we deploy two smartphones, an OBD-II interface and a GPS receiver on the map-construction car. In addition, we mount 16 GSM radios on the roof of the car and a Hall sensor to detect the revolution of wheels in order to get the real travelled distance.

B. Evaluation Methodology

With the prototype system, we apply UPS in a demonstration area of six square kilometers as shown in Figure 6(b). We randomly select 50 locations on roads as reference locations and make marks at these locations. For map construction, we drove the map-construction car to collect all sensory data in this area twice a day for 45 days from May 4th to June 18th, 2014. For comparison, we also construct a GPS map with the collected GPS data following the same method for context map construction. We drove the localization car following arbitrary routes in this area for three days from June 19th to 21st with speed varying from 20km/h to 80km/h, collecting all data for localization performance evaluation. We encountered both heavy and light traffic when those traces were collected.

We conduct extensive trace-driven simulations to evaluate the performance of UPS in urban settings and compare with GPS using the metric of *localization error*, which refers to the absolute distance between the estimated location and the ground truth of a reference location. As reference locations are physically marked, they can be identified in both the map-construction data trace and the localization data trace, and labeled in a constructed fingerprint map. To measure the localization errors, we use the following methods for UPS and GPS, respectively:

- **UPS:** we try to localize each reference location in the localization trace on a constructed map and calculate the distance between the localized location and the labeled location of that reference point on the map.
- **GPS:** As GPS reports may deviate from roads, we first project the GPS reports received at each reference location to a nearest road segment, and calculate the distances between the projected locations and the labeled locations on the GPS map as the localization errors of GPS.

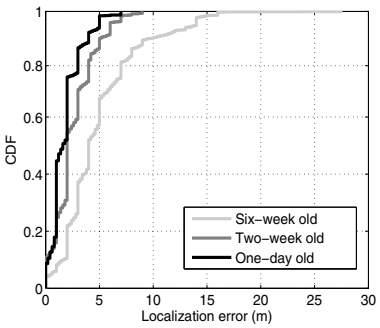


Fig. 7: Localization errors vs. context maps of different ages.

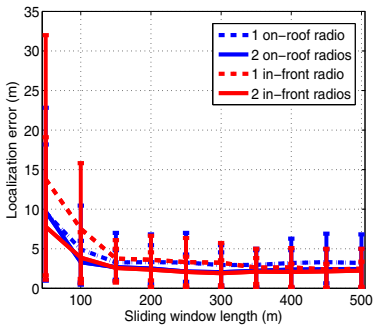


Fig. 8: Localization errors vs. the sliding window length under different radio configurations.

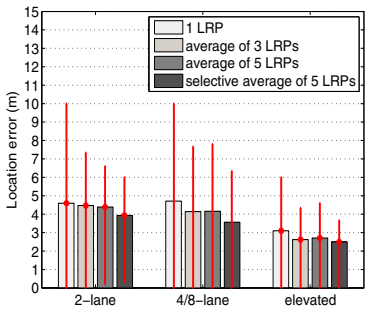


Fig. 9: Localization errors vs. various urban scenarios using variant aggregation schemes.

C. Impact of Fingerprint Evolution

We first examine how context map of cellular-aware trajectories evolve over time. We use trace collected on May 4th, June 4th, and June 18th to construct a context map, respectively. For each map, we use all reference locations in the localization trace collected with one GSM radio placed on the roof of the localization car for testing. For each reference point, the most-recent cellular-aware trajectory of 300 meters is used to conduct the maximum similarity search.

Figure 7 plots the CDFs of localization errors using different maps. It can be seen that, in general, cellular-aware trajectories are very reliable. For example, even using a six-week-old map for localization, the average localization error is about 4.6 meters. There is a tradeoff between the localization performance and the map updating cost.

D. Effect of Sliding Window and Radio Configuration

We study the effect of the sliding window size and radio configuration to the localization performance of UPS, using a two-week-old map constructed on June 4th. We vary the length of the sliding window from 25 meters to 500 meters and vary the number and placement of GSM radios. For each setting, we use all reference locations in the localization trace for testing and calculate the average localization errors.

Figure 8 plots the average and 90% confidence interval of localization errors as a function of the sliding window length under different radio configurations. It can be seen that, as the window length increases, the localization errors decrease and remain stable when the window length is larger than 200 meters. The reason is that the temporal stability goes better when sufficiently long cellular-aware trajectories are used for localization. It can also been seen that adding more GSM radios can help improve the localization performance of UPS and the radio placement has less impact on the performance.

E. Impact of Urban Environment

In this experiment, we examine the impact of urban environment to the localization performance. We divide the localization trace collected with two GSM radios placed in the front of the vehicle according to the four different scenarios, i.e., on 2-lane roads (normal urban roads), on 4- and 8-lane roads (major urban roads) and on roads that are under elevated roads (major urban roads with semi-open condition). For each reference location, we use the most-recent trajectory of 200 meters to conduct online localization algorithm. In addition, for each reference location, we also randomly select three and five trajectories of 200 meters within a 20-meter interval centered at that reference location to localize three and five LRPs. With multiple LRPs, two aggregation schemes, i.e., normal average and selective average (the maximum and minimum estimates are discarded before the rest estimates are averaged), are adopted to retrieve the final localization results.

Figure 9 plots the average and the 90% confidence interval of localization errors under different urban scenarios with various aggregation schemes. It can be seen that, in all scenarios, UPS can achieve extraordinary localization performance. For example, the average localization error is less than five meters. It is also surprising to see that UPS can achieve best localization performance when it is on under elevated roads. The main reason is that the success of UPS relies on the wide availability and rich temporal-spatial features of GSM signals and therefore has less to do with the road type. In addition, using multiple LRPs and the selective aggregation scheme can effectively reduce the localization errors. Thus, we use the selective average over five LRPs to estimate vehicle locations and compare UPS with GPS in the next experiment.

F. Performance Comparison

In this experiment, we compare UPS with GPS and POLS [13] in urban settings. We divide the localization trace collected with two GSM radios placed in the front of the vehicle according to the four different scenarios, i.e., on 2-lane roads (normal urban roads), on 4- and 8-lane roads (major urban roads) and on roads that are under elevated roads (major urban roads with semi-open condition). For each reference location, we use the most-recent trajectory of 200 meters to conduct online localization algorithm. In addition, for each reference location, we also randomly select three and five trajectories of 200 meters within a 20-meter interval centered at that reference location to localize five LRPs. For each reference location, the selective average (the maximum and minimum

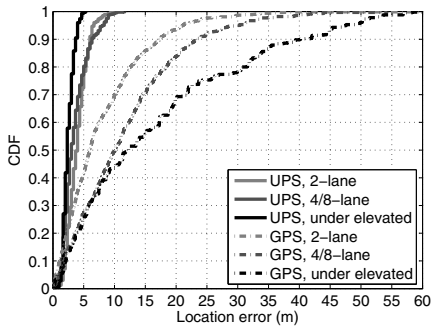


Fig. 10: Comparison between UPS and GPS in urban settings.

estimates are discarded before the rest estimates are averaged) over five LRPs is adopted to localize that reference point.

Figure 10 plots the CDFs of localization errors. It can be seen that UPS is more robust and stable in all kinds of urban scenarios compared with GPS. For example, when tested on 2-lane, 4/8-lane and under-elevated roads, UPS achieves low localization errors of 6, 6.5 and 3.5 meters with a precision of 90%, respectively, whereas GPS has much higher localization errors of 17.5, 23.5 and 40 meters, respectively. When compared with POLS, UPS has an absolute advantage where POLS has an average localization errors about 94 meters over all scenarios. This is mainly because POLS only utilizes the RSSI information of a small number of channels allocated in the current and neighbouring cells and the information collected on single positions are used for localization.

VII. Conclusion and Future Work

In the paper, we show that cellular-aware trajectories have ideal temporal-spatial characteristics as fingerprints for localization. We have proposed an urban localization scheme called UPS which can achieve stable localization performance under variant urban environments and at the same time needs a minimum hardware deployment. We have implemented a prototype system of UPS and verified the efficacy of UPS design through extensive field studies. The experimental results show a promising performance that UPS can achieve a high localization accuracy of 4.2 meters on average and 5.3 meters with a 90% precision. In future, we will further explore other ambient wireless signals such as TV band, 3G and LTE for fingerprinting. Another interesting direction would be how to leverage the power of crowdsourcing to reduce the cost of constructing and updating the context map.

References

[1] Y.-T. Chan, W.-Y. Tsui, H.-C. So, and P.-c. Ching, "Time-of-arrival based localization under nlos conditions," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 17–24, 2006.

[2] K. Lin, A. Kansal, D. Lymberopoulos, and F. Zhao, "Energy-accuracy trade-off for continuous mobile device location," in *Proceedings of ACM MobiSys*, 2010.

[3] J. Paek, J. Kim, and R. Govindan, "Energy-efficient rate-adaptive gps-based positioning for smartphones," in *Proceedings of ACM MobiSys*, 2010.

[4] H. S. Ramos, T. Zhang, J. Liu, N. B. Priyantha, and A. Kansal, "Leap: a low energy assisted gps for trajectory-based services," in *Proceedings of ACM UbiComp*, 2011.

[5] M. B. Kjærgaard, S. Bhattacharya, H. Blunck, and P. Nurmi, "Energy-efficient trajectory tracking for mobile devices," in *Proceedings of ACM MobiSys*, 2011.

[6] E. Elnahrawy, J.-A. Francisco, and R. P. Martin, "Bayesian localization in wireless networks using angle of arrival," in *Proceedings of the 3rd ACM International Conference on Embedded Networked Sensor Systems*, 2005.

[7] M. Youssef, A. Youssef, C. Rieger, U. Shankar, and A. Agrawala, "Pinpoint: An asynchronous time-based location determination system," in *Proceedings of ACM MobiSys*, 2006.

[8] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proceedings of ACM MobiCom*, 2000.

[9] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter *et al.*, "Place lab: Device positioning using radio beacons in the wild," in *Pervasive computing*. Springer, 2005, pp. 116–133.

[10] J. Paek, K.-H. Kim, J. P. Singh, and R. Govindan, "Energy-efficient positioning for smartphones using cell-id sequence matching," in *Proceedings of ACM MobiSys*, 2011.

[11] "Google maps for mobile," http://www.google.com/mobile/maps/.

[12] "Skyhook," http://www.skyhookwireless.com/.

[13] M. Y. Chen, T. Sohn, D. Chmelev, D. Haehnel, J. Hightower, J. Hughes, A. LaMarca, F. Potter, I. Smith, and A. Varshavsky, "Practical metropolitan-scale positioning for gsm phones," in *Proceedings of ACM UbiComp*, 2006.

[14] A. Varshavsky, M. Y. Chen, E. de Lara, J. Froehlich, D. Haehnel, J. Hightower, A. LaMarca, F. Potter, T. Sohn, K. Tang *et al.*, "Are gsm phones the solution for localization?" in *Proceedings of 7th IEEE Workshop on Mobile Computing Systems and Applications*, 2005.

[15] M. Ibrahim and M. Youssef, "Cellsense: An accurate energy-efficient gsm positioning system," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 286–296, 2012.

[16] P. Bahl and V. N. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," in *Proceedings of IEEE INFOCOM*, 2000.

[17] M. Youssef and A. Agrawala, "The horus wlan location determination system," in *Proceedings of ACM MobiSys*, 2005.

[18] Y. Gwon and R. Jain, "Error characteristics and calibration-free techniques for wireless lan-based location estimation," in *Proceedings of the 2nd ACM International Workshop on Mobility Management and Wireless Access Protocols*, 2004.

[19] J. Xiao, K. W. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: passive device-free indoor localization using channel state information," in *Proceedings of IEEE ICDCS*, 2013.

[20] "Osmocombb project," http://bb.osmocom.org/trac/.

[21] G. Chandrasekaran, T. Vu, A. Varshavsky, M. Gruteser, R. P. Martin, J. Yang, and Y. Chen, "Tracking vehicular speed variations by warping mobile phone signal strengths," in *Proceedings of IEEE PerCom*, 2011.

[22] H. Han, J. Yu, H. Zhu, Y. Chen, J. Yang, Y. Zhu, G. Xue, and M. Li, "Senspeed: Sensing driving conditions to estimate vehicle speed in urban environments," in *Proceedings of IEEE INFOCOM*, 2014.