A Mutual Information–Based Infrastructureless Radio Frequency Positioning System via Deep Learning

Alexander I-Chi Lai, Chung-Yuan Chen, and Ruey-Beei Wu

Abstract — We proposed a novel, scalable, spatial positioning solution based on existing radio frequency (RF) signals and available infrastructure by harvesting information from the signals, distilling the readings to get the mutual information gain of each perceived RF characteristic at the corresponding location of interest, and processing such information with a deep neural network for localization. The solution leveraged all available characteristics of perceived RF signal sources instead of filtering them out. We built and tested a scalable positioning prototype on the Google Kubernetes containerization framework accordingly. The preliminary results achieved at least 1.8 m (6 ft) of accuracy within 90% cumulative distribution and 1.08 m (3.3 ft) mean accuracy by solely using 802.11ac received signal strength fingerprints. The prototype performed approximately three times as good as the referenced baseline of weighted k-nearest neighbors.

1. Introduction

Traditionally, positioning has been done by establishing purpose-made localization infrastructures, such as the Global Navigation Satellite System, to range among the interested targets and the “anchors” within the infrastructure for multilateration and/or multianngulation. However, the effectiveness of infrastructure-based approaches is compromised by irregular environmental factors, particularly indoors. They also suffer from high infrastructure costs and are prone to disturbance from natural disasters, human calamities, and technological obsolescence.

Alternatively, the scene analysis approaches require only perceived radio frequency (RF) signals and are virtually free from infrastructure deployment overhead. Since the early 2000s, fingerprinting [7–11], especially that based on Wi-Fi (IEEE 802.11) and received signal strength (RSS), has become a mainstream approach for scene analysis–based localization due to Wi-Fi’s pervasiveness. Various fingerprinting localization approaches, such as k-nearest neighbors [7, 10], weighted k-nearest neighbors (WKNN) [9], and variants with heuristic filters [8], assisted localization [12], and model enhancements of probabilistic analyses [13–15], have therefore evolved.

Information theory–oriented enhancements also go one step further by processing the entropy/information gain of digitized RSS readings [15–19] primarily for better Wi-Fi access point (AP) selection. Another option is to use characteristics such as channel state information (CSI) for fingerprint ingredients. The additional information carried by the finer-grained CSI (though unavailable on many existing Wi-Fi platforms) achieves a better ranging and activity recognition [20, 21] as well as fingerprint-based localization with better accuracy [22–24]. Recently, learning-based localization approaches, especially deep-learning ones like deep neural network (DNN), have rapidly gained attention [25–28].

While fingerprinting bypasses using the signal-emitting infrastructure or “anchors,” it still faces challenges of degraded positioning accuracies and biased estimations [2]. Particularly in indoor cases, the distribution of fingerprints tends to fluctuate [2], indicating that fingerprint information is difficult to process properly. Furthermore, classical scene analyses use only a fraction of the signal information, leading to unsatisfactory accuracy. The manual data collection method is time consuming and subject to human error.

We tackled these challenges by devising a mutual information–based, infrastructureless, scalable positioning solution based on deep learning and enhanced with robotic site-surveying mechanisms. Specifically, various characteristics of perceived RF signal sources at a surveying location, including but not limited to RSS, CSI, link quality indicators, and round-trip time (RTT), were digitized and calibrated by our purpose-built site-surveying robots with advanced light detection and ranging (LiDAR) and simultaneous localization and mapping (SLAM) capabilities. The extracted location readings were collectively treated as the fingerprints.

Information gains of fingerprints and the collective global mutual information of the localization site were calculated to prioritize signal characteristics for further matching. During positioning, the signal readings were fed into a DNN combining one classifier and multiple regressors to match the stored fingerprints and estimate the actual location of the target. A prototype indoor positioning system (IPS) was accord-

Manuscript received 24 August 2020.

This work was financially supported by the “Center for electronics technology integration (NTU-108L900503)” from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Alexander I-Chi Lai, Chung-Yuan Chen, and Ruey-Beei Wu are with the Department of Electrical Engineering, National Taiwan University, No. 1, Section 4, Roosevelt Road, Da’an District, Taipei City, 10617, Taiwan; e-mail: alexiuslai@ntu.edu.tw.
2. Data Processing by Mutual Information

In information and communication theory, information gain is the amount of information obtained about a signal by observing another signal, rendering as a good indicator to measure the scene similarities—and therefore the proximity of locations—in infrastructureless localization. In contrast to entropy-based approaches, the proposed solution explores the relationship of the signal characteristics and the surveying locations instead of filtering out signal sources by AP selection.

Assume that the site survey is conducted at $R$ reference points (RPs) of locations $r_1, \ldots, r_R$ in the target site with $F$ perceivable signal features, including but not limited to RSS, CSI, link quality indicators, and RTT, from certain signal sources, such as Wi-Fi APs. A fingerprint is a collection of the perceivable signal features, including $X$ given reference points $(R)$ of locations $X$. In information and communication theory, information gain (IG) of a feature $X_j$ is obtained as the decrease in entropy of $X_j$ given $R = r$:

$$ IG(X_j, R = r) = H(X_j) - H(X_j|R = r) $$

Also, the mutual information (MI), a global indicator for evaluating the contingency of the $j$th feature $X_j$ and the reference points $R$, is defined as:

$$ MI(X_j, R) = H(X_j) - \sum_{r \in R} P_R(r) \cdot H(X_j|R = r) $$

In the proposed solution, MI and IG were used to prioritize each RP signal feature, matching for better results. The following data processing was achieved:

- Feature categorization: A feature vector $x$ can be decomposed into subvectors $x_{ci}, (c \in C)$. In this study, the two categories used were signal frequencies—2.4-GHz and 5-GHz bands, respectively—to gauge the relative influence level per category.

- Intrasample normalizations: Extract the relative strength of each feature element within feature vectors to suppress device miscalibration errors. Each element $x_{ij}$ of the feature vector $x_i$ undergoes the following normalizations:

$$ x_{ij} := [\min(x_i)]/([\max(x_i) - \min(x_i)] $$

- Feature recomposition: Normalized subvectors were recomposed into vectors of 2.4 GHz only, 5.8 GHz only, or a combination of both. This enabled the localization mechanism to cope with single-band device variations and protocol backward compatibilities.

3. Deep Neural Network for Positioning

Distilled and normalized fingerprint samples were fed into our positioning DNN for further localization. Different from classifier- or regressor-oriented deep-learning approaches [25–28], the primary DNN in the study, inspired from the YOLO unified detection model [29], structured a multilabel classifier and multiple coordinate regressors with a final aggregation layer. On getting a distilled and normalized input feature vector from the localization target, the DNN classifier determined the probability that a localization target appeared within the maximal influential range $(\gamma)$ of an RP. Meanwhile, the regressor of that RP estimated the normalized offset vector (distance and direction) between the target and the RP. The distilled input feature vector with the $R$ pairs of the probability and offset vector combination was fed into three fully connected hidden layers to obtain the probabilistic mapping from the input vector to the RPs. It was then converted into the final coordinate estimation of the location target in the exiting aggregation stage.

Following the process, each RP’s feature information is preserved rather than filtered out. The estimation error of each regressor is confined by the maximal influential range $\gamma$ of each RP instead of being scaled by the span of the target area. The error of the final estimation is reduced by spatial diversity when
combining estimations from multiple RPs. Refer to Figure 1 for the design and configuration of the primary DNN we designed.

### 4. Prototype Design and Implementation

For comparative testing, we constructed an IPS prototype of our mutual information–based DNN. Elastic cloud computing services were employed to meet computing power demand. The scalable back end was software container based and could accumulate fingerprint data from Wi-Fi APs with a large set of RPs. Our prototype, built on the scalable Google Kubernetes containerization framework, as shown in Figure 2, was operational on a three-node, 10-Gb Ethernet-connected private Linux cluster with dual 64-bit Intel Xeon 8-core processors, 64 GB RAM, 6 TB storage, and one NVIDIA Quadro P1000 graphics processing unit on each server node. The DNN module was developed in Python and Keras 2.3.0 with Tensorflow 1.14.0. A Web-based visualizer was created to show the Wi-Fi station heat map, signal strength, occurrence frequencies, receiving points, information gains, and mutual information for data analysis.

Our custom-built site-surveying robots, as shown in Figure 3, navigated through the RPs within the testing site and gathered authenticated fingerprints for improving spatial models. ROBOTIS TurtleBot3–based robots were equipped with a seventh-generation Intel Core i5 processor, multiple 802.11ac Wi-Fi adapters, an Intel RealSense camera, and enhanced LiDAR and attended SLAM capability. The SLAM data were combined with the building information management floor map and calibrated using a GIMP raster graphics editor to achieve a global static map for odometry-LiDAR adaptive Monte Carlo localization real-time positioning, as shown in Figure 3. Extracted raw RF data were uploaded into a MongoDB server at the IPS back end for further processing and DNN training. The robots significantly increased the surveying efficiency and virtually eliminated human measurement errors.

### 5. Experimental Results

We conducted a series of validation experiments on the prototype using the Ming-Da Building at National Taiwan University, a seven-story, 12,500-m² reinforced-concrete building. In addition to existing Wi-Fi deployment, 32 ultra-wideband transponders and 13 additional Wi-Fi APs supporting IEEE 802.11mc (i.e., Wi-Fi RTT) were added to calibrate and evaluate performances of different positioning methodologies. We tested our prototype in a 30 m × 12 m area at level 5F of the Ming-Da Building. The distribution of the 349 RPs in the surveyed area was illustrated in Figure 4. Of the more than 1,500 discovered Wi-Fi APs, 191 were selected by the analyses in [19] to construct the fingerprint data set. Refer to Table 1 for the data set used in the study.
Table 2 shows the comparisons of our DNN against the classical WKNN and a “naive” DNN with a single coordinate estimator positioning multiple Raspberry Pis and Google Pixel 4 phones. The results showed that our DNN-enabled prototype can achieve 1.08 m (3.3 ft) mean accuracy and no worse than 1.86 m (6 ft) accuracy within 90% cumulative distribution by solely using 802.11ac RSS fingerprints due to limited availability of CSI or other properties. That is approximately three times better than the referenced WKNN baseline. When Wi-Fi RTT information was incorporated, 1.2 m (4 ft) accuracy within a 90% cumulative distribution was achieved (not shown in Table 2). Refer to Figure 5 for the testing results of our mutual information–enhanced DNN on the probability distribution function (PDF), the cumulative distribution function (CDF), and the positioning/visualization of localization error estimations of all samples. Note that the positioning results are also insensitive to the disturbance of any specific AP, such as being off or moved, or in non–line-of-sight conditions, due to the massive amount of effective radio map data accumulated.

6. Conclusion

The results validated the effectiveness of combining mutual information, probabilistic classifiers, and coordinate regressors and demonstrated the advantages of using fingerprint detail rather than filtering out the signal sources. The same design can accommodate heterogeneous fingerprint features to achieve better accuracy. Scalable positioning back ends are portable to other platforms, providing adequate computing power for setups of any size. Finally, site-surveying robots significantly improve infrastructureless localization feasibilities.

7. References


