CSI-based Indoor Localization

Kaishun Wu *Member, IEEE* Jiang Xiao, *Student Member, IEEE,* Youwen Yi *Student Member, IEEE,* Dihu Chen , Xiaonan Luo and Lionel M. Ni *Fellow, IEEE*

Abstract—Indoor positioning systems have received increasing attention for supporting location-based services in indoor environments. WiFi-based indoor localization has been attractive due to its open access and low cost properties. However, the distance estimation based on received signal strength indicator (RSSI) is easily affected by the temporal and spatial variance due to the multipath effect, which contributes to most of the estimation errors in current systems. In this work, we analyze this effect across the physical layer and account for the undesirable RSSI readings being reported. We explore the frequency diversity of the subcarriers in OFDM systems and propose a novel approach called FILA, which leverages the channel state information (CSI) to build a propagation model and a fingerprinting system at the receiver. We implement the FILA system on commercial 802.11 NICs, and then evaluate its performance in different typical indoor scenarios. The experimental results show that the accuracy and latency of distance calculation can be significantly enhanced by using CSI. Moreover, FILA can significantly improve the localization accuracy compared with the corresponding RSSI approach.

Index Terms—Indoor localization, Channel State Information, RSSI, Physical Layer.

1 Introduction

Localization is one of the essential modules of many mobile wireless applications. Although Global Positioning System (GPS) works extremely well for an open-air localization, it does not perform effectively in indoor environments due to the disability of GPS signals to penetrate in-building materials. Therefore, precise indoor localization is still a critical missing component and has been gaining growing interest from a wide range of applications, e.g., location detection of assets in a warehouse, patient tracking inside the building of the hospital, and emergency personnel positioning in a disaster area.

A great number of researches have been done to address the indoor localization problem. Many range-based localization protocols compute positions based on received signal strength indicator (RSSI), which represents the received power level at the receiver. According to propagation loss model [1], received signal power monotonically decreases with increasing distance from the source, which is the foundation of the model-based localization. Most of the existing radio frequency (RF)-based indoor localization are based on the RSSI values [1]–[5]. More related work is in the supplemental file. However, we claim that the fundamental reasons why RSSI is not suitable for indoor localization are from two aspects: First, RSSI is measured from the RF signal at a per packet level, which is difficult to obtain an accurate value. According to our measurement in a typical indoor environment as shown in Fig. 1, the variance of RSSIs collected from an immobile receiver in one minute is up to 5dB. Second, RSSI is easily varied by the multipath effect. In theory, it is possible to establish a model to estimate the separation distance using the received power. In reality, however, the propagation of a RF wave is attenuated by reflection when it hits the surface of an obstacle. In addition to the line-of-sight (LOS) signal, there are possibly multiple signals arriving at the receiver through different paths. This multipath effect is even more severe in indoor environments where a ceiling, floor and walls are present. As a result, it is possible for a closer receiver to have a lower RSSI than a more distant one. Consequently, a simple relationship between received power and separating distance cannot be established. Therefore, this time-varying and vulnerable RSSI value creates undesirable localization errors.

We argue that a reliable metric provided by commercial NICs to improve the accuracy of indoor localization is in need. Such metric should be more temporal stable and provide the capability to benefit from the multipath effect. In current widely used Orthogonal Frequency Division Multiplexing (OFDM) systems, where data are modulated on multiple subcarriers in different frequencies and transmitted simultaneously, we have a value that estimates the channel in each subcarrier called Channel State Information (CSI). Different from RSSI, CSI is a fine-grained value from the PHY layer which describes the amplitude and phase on each subcarrier in the frequency domain. In contrast to having only one

- *K. Wu is with National Engineering Research Center of Digital Life, State-Province Joint Laboratory of Digital Home Interactive Applications, Sun Yat-sen University, Guangzhou, China. He is also affiliated at Fok Ying Tung Graduate School, HKUST. Email:{kwinson}@ust.hk.*
- *Jiang Xiao, Youwen Yi and Lionel M. Ni are with Department of Computer Science and Engineering, HKUST. E-mail: {jxiao,ywyi,ni}@cse.ust.hk.*
- *Dihu Chen is with School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China. E-mail: {stscdh}@mail.sysu.edu.cn.*
- *Xiaonan Luo is with National Engineering Research Center of Digital Life, State-Province Joint Laboratory of Digital Home Interactive Applications, Sun Yat-sen University, Guangzhou, China. E-mail: {lnslxn}@mail.sysu.edu.cn.*

RSSI per packet, we can obtain multiple CSIs at one time. More importantly, the CSIs over multi-subcarriers will travel along different fading or scattering paths on account of the multipath effects. It then naturally brings in the frequency diversity attribute of CSI, which each subcarrier has different amplitudes and phases. By exploiting the frequency diversity, we can construct a unique "fingerprinting" indicating each location on the radio map. According to these advantages, it is favorable to leverage the CSI to improve the performance of indoor location fingerprinting. And thus, designing a precise tracking/localization system becomes possible.

Based on CSI, in this paper, we present the design and implementation of FILA, a novel cross-layer approach based on OFDM for indoor localization using WLANs.

In summary, the main contributions of this paper are as follows.

1) We design FILA, a cross layer approach that enables fine-grained indoor localization in WLANs. FILA includes two parts, the first one is CSI-based propagation model and the second one is CSI-based fingerprinting. To the best of our knowledge, FILA is the first to use fine-gained PHY layer information (CSI) in OFDM to build a propagation model so as to improve indoor localization performance. And it is also the first time to take advance of the combination of the fine-grained PHY layer information CSI with frequency diversity and multiple antennas with spatial diversity for indoor location fingerprinting.
2) We implement FILA in commercial 802.11 NICs and conduct extensive experiments in several typical indoor environments to show the feasibility of our design.
3) Experimental results demonstrate that FILA significantly improves the localization accuracy as compared to the corresponding traditional RSSI-based approach.

The rest of this paper is organized as follows. In Section. 2, we introduce some preliminaries. This is followed by the system architecture design in Section 3. In Section 4, we demonstrate the CSI-based propagation model. In Section 5, we illustrate the methodology of CSI-based fingerprinting. The implementation of FILA and experimental evaluations are presented in Section 6. Finally, conclusions are presented and suggestions are made for future research in Section 7.

2 Preliminaries

In this section, we introduce the CSI value which is the foundation of FILA design. And the background information of the OFDM system can be found in supplemental file.

2.1 Channel State Information

Based on OFDM, channel measurement at the subcarrier level becomes available. Nowadays, adaptive transmission systems in wireless communication always improve the throughput by utilizing some knowledge of the channel state to adapt or allocate transmitter resources [6].

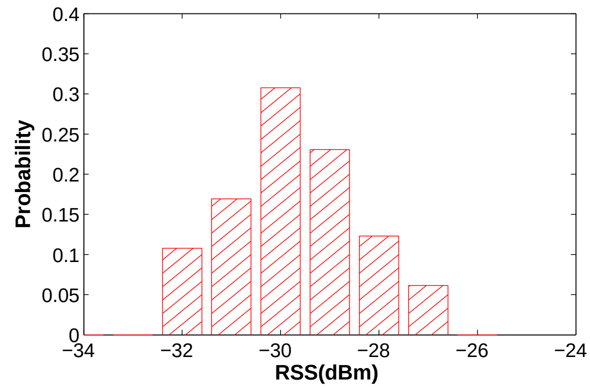


Fig. 1: Temporal variance of RSSI.

Channel state information or channel status information (CSI) is information that estimates the channel by representing the channel properties of a communication link. More specifically, CSI describes how a signal propagates from the transmitter(s) to the receiver(s) and reveals the combined effect of, for instance, scattering, fading, and power decay with distance. In summary, the accuracy of CSI greatly influences the overall OFDM system performance. It is worth pointing out that according to the definition of CSI, only OFDM-based WLAN systems can demonstrate the frequency diversity in CSI since they use multiple subcarriers for data transmission. In another word, other modulation schemes like DSSS cannot provide this value.

In a narrowband flat-fading channel, the OFDM system in the frequency domain is modeled as

$$y = Hx + n, \tag{1}$$

where y and x are the received and transmitted vectors, respectively, and H and n are the channel matrix and the additive white Gaussian noise (AWGN) vector, respectively.

Thus, CSI of all subcarriers can be estimated according to (1) as

$$\hat{H} = \frac{y}{x}, \tag{2}$$

which is a fine-grained value from the PHY layer that describes the channel gain from TX baseband to RX baseband.

3 System Design

In this section, we first give an overview of the system architecture. Challenges in the system design are presented in the supplemental file.

3.1 System Architecture

FILA system is built based on the current communication system and thus compatible to the under layer design. More precisely, no modification is required at

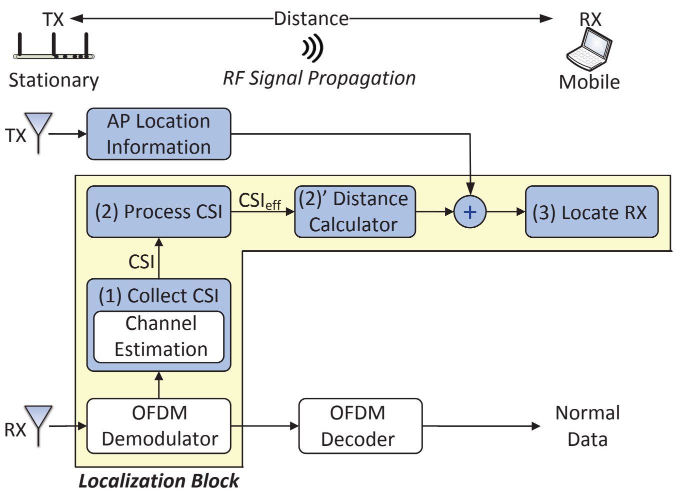


Fig. 2: System Architecture.

the transmitter end (TX–the AP), while only one new component for CSI processing is introduced for localization purposes at the receiver end (RX–the target mobile device). Fig. 2 demonstrates the detailed design of the system architecture. For traditional packet transmission in wireless communication, only the demodulated signal is exported to the decoder for message content retrieval. However, a prerequisite in FILA localization system is that it should be able to export the CSI value after the normal demodulation process. Such that we devise a localization block to exploit the CSI information.

In our designated localization block, the CSI collected from 30 groups different subcarriers will firstly be processed. After running the proposed algorithm, we can obtain the effective CSI in an efficient time constraint. Then the effective CSI will be used to estimate the location of the target object. As mentioned in the previous section, CSI value is the channel matrix from RX baseband to TX baseband which is needed for channel equalization. Therefore, there is no extra processing overhead when obtaining the CSI information. Nevertheless, RSSI is obtained at the receiver antenna in the 2.4 GHz radio frequency before down convert to the IF and baseband. Therefore, the free space model that built for RSSI-based localization approaches can't be directly applied to process the CSI value. We need to refine such radio propagation model according to the CSI information and compute the distance based on the proposed one. Finally, as the AP location information is obtained from the network layer while CSI is collected from the physical layer, we then use the simplest trilateration method to obtain the location. For the fingerprinting, we leverage the CSI values including different amplitudes and phases at multiple propagation paths, known as the frequency diversity, to uniquely manifest a location. We then present a probability algorithm with a correlation filter to map object to the fingerprints.

In our FILA system architecture design, CSI is only processed by a newly designed localization block if needed. Owing to the fact, FILA can be applied concurrently with the original packet transmission. In other words, it will not introduce additional overhead during the data transmission.

4 CSI-BASED PROPAGATION MODEL

In this section, we leverage the fine grained CSI value instead of RSSI to build a propagation model and address the indoor localization issue. The CSI-based propagation model can be built based on three following steps.

1) CSI Processing: First, we need to mitigate estimation error by effectively processing the CSI value denoted as CSI_{eff}. This is known as the prerequisite of the ongoing two steps.
2) Calibration: Afterwards, we develop a refined indoor propagation model and a fast training algorithm to derive the relationship between CSI_{eff} and distance.
3) Location Determination: By receiving the APs coordinates in network layer and CSI_{eff} values from physical layer, we apply the revised propagation model and trilateration method to accomplish the localization. This part is in the supplemental file.

4.1 CSI Processing

For wireless communication, attenuation of signal strength through a mobile radio channel is caused by three nearly independent factors: path loss, multipath fading, and shadowing. The path loss characterizes the property that the signal strength decays as the distance between the transmitter and receiver increases, which is the foundation of our CSI-based localization. Multipath fading is a rapid fluctuation of the complex envelope of received signal caused by reception of multiple copies of a transmitted signal through multipath propagation. Shadowing represents a slow variation in a received signal strength due to the obstacles in propagation path. Therefore, before establishing the relationship between CSI and distance, we need to mitigate the estimation error introduced by multipath fading and shadowing.

4.1.1 *Time-domain Multipath Mitigation*

The first concentration of our design is that the system must be capable of dealing with the challenge of operating over a multipath propagation channel. Since multipath effect will introduce Inter-Symbol-Interference (ISI), cyclic prefix (CP) is added to each symbol to combat the time delay in OFDM systems. However, the CP technique is helpless for the multiple reflections within a symbol time.

For narrow-band systems, these reflections will not be resolvable by the receiver when the bandwidth is less than the coherence bandwidth of the channel. Fortunately, the bandwidth of 802.11n waveforms is 20MHz (with channel bonding, the bandwidth could be 40MHz), which provides the capability of the receiver to resolve the different reflections in the channel. We propose a multipath mitigation mechanism that can distinguish the LOS signal or the most closed NLOS from other reflections in the expectation of reducing the distance estimation error.

The commonly used profile of multipath channel in the time domain is described as follow,

$$\mathrm{h}(\tau) = \sum_{k=0}^{\mathrm{L}_p - 1} \alpha_k \delta(\tau - \tau_k), \tag{3}$$

where L_p is the number of multipath channel component. α_k and τ_k are the amplitude and propagation delay of the k-th path.

In practice, OFDM technologies are efficiently implemented using a combination of fast Fourier Transform (FFT) and inverse fast Fourier Transform (IFFT) blocks. The 30 groups of CSI represent the channel response in frequency domain, which is about one group per two subcarriers. With IFFT processing of the CSI, we can obtain the channel response in the time domain, i.e., $h(t)$. From Fig. 3 we can observe that the LOS signal and multipath reflections come with different time delay, and generally the LOS signal has higher channel gain, so we can use a trunk window with the first largest channel gain in the center to filter out those reflections. If LOS doesn't exist, we can identify the shortest path NLOS reflection. According to Nyquist sampling theorem, wider spectrum leads to higher resolution in the time domain. Due to the bandwidth limitation of WLAN, we can't distinguish all the reflections but we can use this method to reduce the variance induced by multipath effects. The time duration of the first cluster is determined by setting the truncation threshold as 50% of the first peak value. In doing so, we expect to mitigate the estimation error introduced by multipath reflection.

After the time domain signal processing, we reobtain the CSI using FFT. Fig. 3 shows the CSI results after time domain filtering. Note that, commercial NICs embeds hardware circuits for the FFT and IFFT processing, our algorithm introduces ignorable latency to the whole localization procedure.

4.1.2 Frequency-domain Fading Compensation

Moreover, since CSI represents the channel responses of multiple subcarriers, a combination scheme is also introduced to process the CSI value in our system for compensation of the fading of received signals in frequency domain to enhance location accuracy.

In general, when the space between two subcarriers is larger than the coherence bandwidth, they are fading independently. Since the channel bandwidth of 802.11n system is larger than the coherence bandwidth in typical indoor environment, the fading across all subcarriers are frequency-selective. To combat such frequency selectively fading of wireless signals, multiple uncorrelated fading subchannels (multiple frequency subcarriers), that is 30 groups of CSI values are combined at the receiver. Motivation for leveraging the frequency diversity stems from the fact that the probability of simultaneous deep fading occurring on multiple uncorrelated fading envelopes (in our case, resulting from frequency diversity) is much lower than the probability of a deep fade occurring on a single frequency system. Thus, exploiting the wide bandwidth of WLAN that assures sufficiently uncorrelated subcarriers, will reduce the variance in CSIs owing to small scale factors, which appears to be one of the major sources of location determination error. In our FILA system, we weighted average the 30 groups CSIs in frequency domain so as to obtain the effective CSI, which exploits the frequency diversity to compensate the small-scale fading effect.

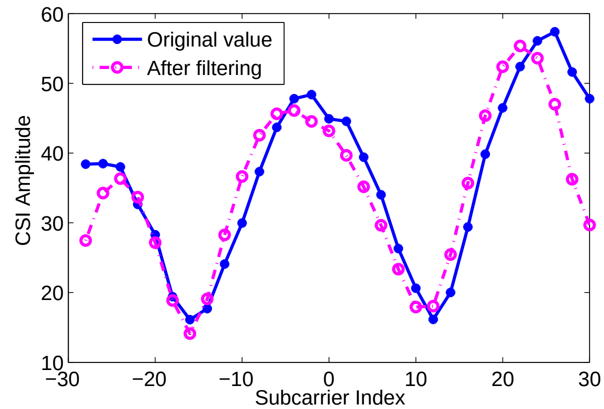


Fig. 3: Time Domain Channel Response.

Given a packet with 30 groups of subcarriers, the effective CSI of this packet is calculated as

$$\mathrm{CSI}_{eff} = \frac{1}{K} \sum_{k=1}^{K} \frac{f_k}{f_0} \times |H_k|, \ \ k \in (-15, 15), \tag{4}$$

where f_0 is the central frequency, f_k is the frequency of the k-th subcarrier, and $|H_k|$ is the amplitude of the k-th subcarrier CSI.

Note that selection of weighting factors are based on the fact that the radio propagation is frequency-related. According to the free space model, the received signal strength is related to the frequency the signal is transmitted. So by this weighting method, we transfer the channel gain from multiple subcarriers to a single subcarrier, i.e., the central one. Next, we will establish the relationship between the CSI_{eff} and distance.

4.2 Calibration

Since CSI value is obtained from the baseband on the receiver side, the radio propagation model [7] for RSSI is no longer suitable for our design. So we develop a refined indoor propagation model to represent the relationship between the CSI_{eff} and distance by revising the free space path loss propagation model, given by:

$$d = \frac{1}{4\pi}\left[\left(\frac{c}{f_0 \times |\mathrm{CSI}_{eff}|}\right)^2 \times \sigma\right]^{\frac{1}{n}}, \tag{5}$$

where c is the wave velocity, σ is the environment factor and n is the path loss fading exponent. Both of the two parameters are dependent on distinctive indoor environments. The environment factor σ represents the gain of the baseband to the RF band at the transmitter side, inversely, the gain of RF band to baseband at the receiver side, and the antenna gains as well. Moreover, for NLOS AP, the σ also includes the power loss due to wall penetration or the shadowing. The path loss fading exponent n is varying depending on the environment. For instance, when the RF signal is propagating along a free space like corridor, the path loss fading exponent n will be around 2. In other cases, such as an office that represents a complex indoor scenario, the exponent could be larger than 4. In an indoor radio channel with clutter in medium, where often the LOS path is augmented with the multipath NLOS at the receiver, signal power decreases with a pass loss fading exponent higher than 2 and is typically in the order of 2 to 4 [8]. Hence, it is not trivial to determine the received signal power and we need to refine the free space propagation model that obeys the analytical and empirical methods. A widely used simplification is to assume that all the path loss exponents that model propagations between the specific receiver and all the APs are equal. This simplification in a typical indoor environment is an oversimplification, since the channel propagation is usually very different depending on the relative position of the mobile client with regard to each AP. Therefore, we calibrate both environment factor σ and the path loss fading exponent n in a per-AP manner.

We propose a simple fast training algorithm based on supervised learning to retrieve the parameters with three anchors in offline phase. In the first step, CSIs of multiple packets are collected at two of the anchors to train the environment factor σ and the path loss fading exponent n for the refined indoor propagation model. In the second step, CSIs collected at the third anchor are used to test the efficiency of the parameter estimation. The two steps run iteratively until convergence. The experimental results in the next section show that this simple algorithm can achieve satisfactory accuracy, more sophisticated training method will be able to obtain better performance.

5 CSI-BASED FINGERPRINTING

In this section, we introduce the methodology of CSI-based location fingerprinting approach. We start by using a mobile device equipped with 802.11 NICs to receive the beacon message from nearby APs at each sample position. The message contains CSI that represents the channel response of multiple subcarriers. We modify the driver and divide the CSIs into 30 groups. Hence, $N = 30$ groups CSI values are collected simultaneously at the receiver that represented as

$$\mathrm{H} = [H_1, H_2, \cdots, H_i, \cdots, H_N]^T,\ i \in [1, 30], \tag{6}$$

where each subcarrier H_i is defined as

$$H_i = |H_i| e^{j \sin\{\angle H_i\}}, \tag{7}$$

where $|H_i|$ is the amplitude response and $\angle H$ is the phrase response of the i_{th} subcarrier.

Then, it comes to CSI processing which is the prerequisite of calculating position likelihood in the positioning phase. We quantify the power of a package, denoted as summational CSI, by adding up the power with respect to 30 groups of subcarriers. Specifically,

$$\mathrm{H}_e = \sum_{i=1}^{I} |H_i|^2,\ i \in [1, 30], \tag{8}$$

5.1 CSI-based Fingerprinting Generation

As the foundation of fingerprinting approach, the measured CSI values are processed to construct a radio map. Since most of the RF-based fingerprint methods consider two spatial dimensions for localization [9], we also follow the principle. Therefore, the two-dimension physical space coordinate of a sample position l_j is $l_j = (l_{j,x}, l_{j,y})$. To generate a radio map, we first extract the statistic determine the number of detectable APs for a sample position. At each reference point, we will estimate the signal strength distribution for each access point at each location. In our location system, the signal strength over all the subcarriers is represented by H_e.

Moreover, another component of the radio map will be normalized CSI values at each subcarrier for each AP. The motivation of leveraging the frequency diversity stems from the fact that CSI benefits from the multipath effects, because the received signal at different positions will be the combination of different reflections. Therefore, these normalized CSI values will reflect the combination result ignoring the large scale fading.

5.2 Positioning Phase

For object location estimation, the target is required to be accurately mapped to the radio map.

Previous works show that the probabilistic approaches such as maximum likelihood provide more accurate results than deterministic ones do in indoor environments [9]. Therefore, we adapt the probability model in [10] except that we use H_e instead of RSS value. Similarly, we treat H_e observed from the AP to the receiver at a fixed location as a Gaussian variable. In the proposed system, we will select K best APs to calculate the probability of the MS at each reference point. The criteria for the best AP selection is that those APs with highest H_e values, because they are more reliable. In our experiment, we fix the K to be 3.

The selected K H_e values obtained by the terminal to be located form a vector $\mathbb{H}_e = [\mathrm{H}_{e,1}, \cdots, \mathrm{H}_{e,K}]$. Then, the position estimation problem is equivalent to finding the l that maximizes the posteriori probability $P(l_j|\mathbb{H}_e)$. According to Bayes' law,

$$P(l_j|\mathbb{H}_e) = \frac{P(l_j)P(\mathbb{H}_e|l_j)}{\sum_i P(l_i)P(\mathbb{H}_e|l_i)} = \frac{P(l_j)P(\mathbb{H}_e|l_j)}{P(\mathbb{H}_e)} \quad (9)$$

Note that $P(l_j)$ is the prior probability that the terminal located at the reference point l_i. In [9], [10], uniform distribution is assumed. In contrast, we will leverage the spatial correlation of the CSI to determine the $P(l_j)$.

Recall that CSI is a fine-grained information, we can observe channel response over multiple subcarriers represented by $\mathrm{H}^k = [H_1, H_2, \cdots, H_i, \cdots, H_N]^T$ for the kth AP. We denote the observed CSI with normalization for each AP as $\mathrm{H}(O) \in \mathbb{C}^{N\times K}$, and the CSI recorded in the radio map for the same set of APs at position l_j as $\mathrm{H}(l_j)$. To quantify the similarity of the observed CSI and the stored "fingerprints" for all the APs, we use the Pearson correlation between them which is defined as

$$\rho_{\mathrm{H}(O),\mathrm{H}(l_j)} = \prod_{k=1}^{K} \frac{cov(\mathrm{H}^k(O), \mathrm{H}^k(l_j))}{\sigma_{\mathrm{H}^k(O)}\sigma_{\mathrm{H}^k(l_j)}}, \quad (10)$$

where each AP is considered to be independent. According to the measurement, the spatial channel correlation will decrease as the distance between the two receiver increases. Therefore, with higher ρ, the position of the terminal will be closer to the reference point. Then, the probability of the terminal on each candidate point is defined as

$$P(l_j) = \frac{\rho_{\mathrm{H}(O),\mathrm{H}(l_j)}}{\sum_{i=1}^{J} \rho_{\mathrm{H}(O),\mathrm{H}(l_i)}} \quad (11)$$

where J is the size of the candidate reference points set.

Considering uncorrelated property between each AP, the likelihood $P(\mathbb{H}_e|l_j)$ can be calculated as,

$$P(\mathbb{H}_e|l_j) = \prod_{k=1}^{K} P(\mathrm{H}_{e,k}|l_j), \quad (12)$$

Since the signal strength at each reference point is modeled as a Gaussian variable which requires less samplings than the histogram approach [11]. At the offline phase, we can obtain the expectation $\bar{\mathrm{H}}_{e,k}$ and variance $\sigma_{e,k}$ corresponding to the $\mathrm{H}_{e,k}$, and the $P(\mathrm{H}_{e,k}|l_j)$ is obtained as

$$P(\mathrm{H}_{e,k}|l_j) = \frac{1}{\sqrt{2\pi\sigma_{e,k}}} \exp \frac{-(\mathrm{H}_{e,k} - \bar{\mathrm{H}}_{e,k})^2}{2\sigma_{e,k}}. \quad (13)$$

The location estimation of the terminal is the weighted average over the whole candidate set,

$$\hat{l} = \sum_{j}^{J} \bar{P}(l_j|\mathbb{H}_e)l_j \quad (14)$$

For the fingerprinting method, the terminal can process CSI by itself and then check the globe or local database for localization. It doesn't need to rely on one pubic server.

The performance of the proposed fingerprinting methodology is evaluated in the following section.

6 Experimental Results

In this section, we present the implementation and experimental evaluation of FILA. First, we describe the experimental setup which can be found in the supplemental file. Then we illustrate the validation results for our refined propagation model. Finally, we evaluate the performance of CSI-based propagation model and its fingerprinting. In our evaluation, we use the performance of corresponding RSSI-based approach based on radio propagation model and trilateration as baseline .

II. Experimental Scenarios

We conduct experiments to show the performance and robustness of our FILA system in four different scenarios in the campus of Hong Kong University of Science and Technology as follows:

1) **Chamber** First, we set up a testbed in a $3m \times 4m$ Chamber to collect the RSSI and CSI as shown in Fig. 4. In general, Chamber is an enclosure that used as environmental conditions for conducting testing of specimen. In our experiment, chamber represents the ideal free space indoor environment which means only LOS signal exists without other multipath reflection or external interference.
2) **Research Laboratory** Then, we deployed FILA in an identical indoor scenario – a $5m \times 8m$ research laboratory as shown in Fig. 5. In the laboratory region, we place three APs on the top of three shelters in three dimensions. The experiment was conducted on a weekday afternoon when there were a couple of students sitting or walking around, which will show the robustness of our system to temporal dynamics of the environment. The laptop was placed at a fixed position at the beginning of the experiment and then moved to the next point along a pre-measured path.
3) **Lecture Theatre** In addition, we chose a larger lecture hall to conduct the localization experiments, which is a $20m \times 20m$ lecture theatre. Since the space is relatively large, the influence of the room size can be explored.
4) **Corridor** Finally, we performed experiments in a corridor environment with multiple offices aside in our academic building, which is $32.5m \times 10m$ covering corridors, rooms and cubicles. In this scenario, we expect to illustrate the impact of the absence of LOS APs on the location accuracy.

6.1 Validate the Refined Model

As the target for precise indoor localization, two most important metrics are used to testify FILA: the accuracy

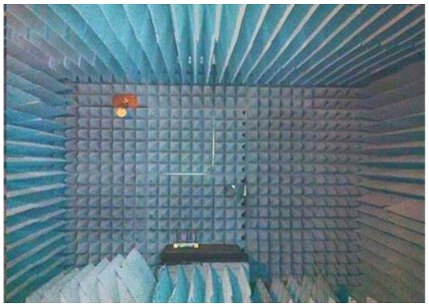


Fig. 4: Chamber

Fig. 5: Research Laboratory

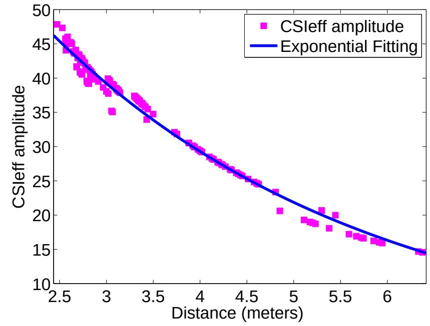


Fig. 6: Relation between CSI_{eff} and Distance.

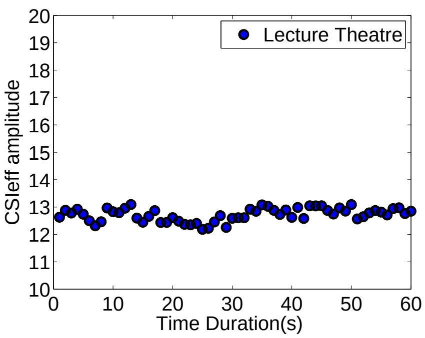


Fig. 7: CSI_{eff} on Temporal Variance

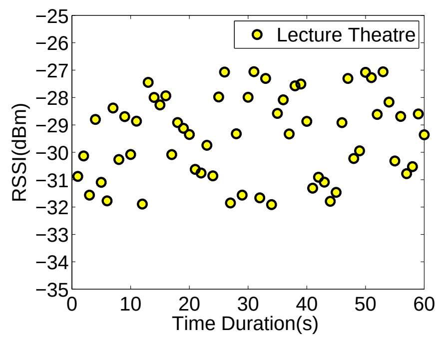


Fig. 8: RSSI on Temporal Variance

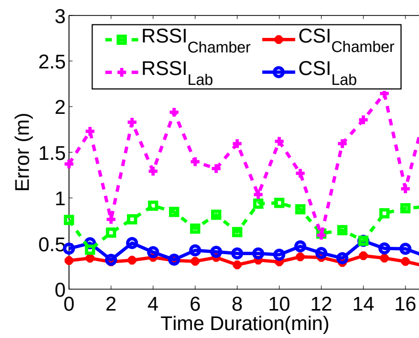


Fig. 9: CSI vs. RSSI on distance estimation in different environments.

and the temporal stability of location estimation. Afterwards, we compared the performance of our CSI-based localization system with the corresponding RSSI-based approach.

6.1.1 Robustness of the Refined Model

One essential aspect that needs to be determined before the localization experiments is whether CSI value can build a relationship with distance. In general, indoor RF signal strength is a non-monotonic function with distance due to multipath and shadowing effects. Fig. 6 illustrates the CSI value approximated by a power function of distance according to our refined propagation model. In diverse scenarios with corresponding environment factor σ, the path loss fading exponent n varies in a range of $[2, 4]$. It is shown that our refined model properly fits the relationship between CSI_{eff} and distance.

6.1.2 Temporal Stability of CSI

Temporal stability is a fundamental criteria in validating the robustness of the localization systems. We thus set out to examine the stability of the proposed new metric CSI_{eff} and RSSI value in time series. It is well-known that RSSI is a fickle measurement of the channel gain because of its coarse packet-level estimation and easily varied by multipath effect. As CSI is fine-grained PHY layer information that provides detailed channel state information in subcarrier level, it is of great importance to figure out whether it will remain in a stable manner in practical environment.

Fig. 7 and Fig. 8 illustrate both the interactions of CSI_{eff} and RSSI values on temporal variance, respectively. It is shown that the received signal power calculated by RSSI has a variance up to 5 dBm, while the variance of CSI_{eff} is within 1 dBm. Therefore, the proposed new matric is much more stable in time domain compared with RSSI.

In Fig. 9, we further investigate the CSI_{eff} value and RSSI value in the chamber and research laboratory so as to discover the effect of any temporal instability on distance estimation. Chamber provides a free space-like environment as it uses specific material that can absorb the non-LOS signals. Thus, multipath effect can be eliminated in chamber environment. However, the result from our experiment shows that even in chamber the RSSI is also varied significantly from time to time due to the inaccurate measurement. In contrast, research lab is a typical multipath environment. Both the static obstacles and dynamic walking around individuals exert the influence on multipath and bring in more intense path loss. In this way, the variance of RSSI becomes even larger and the performance of distance estimation is even worse.

Fig. 7, Fig. 8 and Fig. 9 lead to an essential conclusion: in comparison with RSSI, CSI is more temporally stable in different environments and helps maintain the performance over time. Therefore, FILA can achieve accurate location more quickly than the RSSI-based scheme, which is very crucial for some location-based application

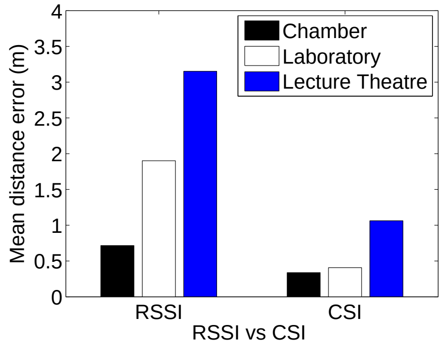


Fig. 10: Mean distance error.

like search and rescue.

6.2 Performance Evaluation of CSI-based Propagation Model

As a target for precise and fast indoor localization, two most important metrics are used to testify this model: the accuracy and the latency of localization, afterwards, we compared the performance of our CSI-based localization systemwith the corresponding RSSI-based approach by using the propagation model.

6.2.1 Accuracy

Accuracy over a Single Link

As the premise of indoor localization, we first investigated the distance determination accuracy of FILA compared with the corresponding RSSI-based approach. The primary source of error in indoor localization is multipath propagation caused by multiple reflections that overlap with the direct LOS subcarrier at the receiver. FILA takes advantage of the fine-grained trait to mitigate such multipath effect, and exploits the frequency diversity to compensate the frequency-selective shading. We repeated the distance measurement experiments across 10 different locations in chamber, research laboratory and lecture theatre, respectively. For some positions with serious multipath effect, FILA achieves up to 10 times accuracy gain over the corresponding RSSI-based scheme. Fig. 10 illustrates the mean distance errors in three different environments. Our evaluation shows that FILA can outperform the corresponding RSSI-based scheme by around 3 times for the distance determination of a single link.

To assess the effectiveness of the CSI-based localization approach, in the following we evaluate the accuracy of FILA in different typical indoor environments.

Localization Accuracy in Single Room

In the experiments conducted in the research laboratory, we fix three APs on the top of the shelters. The mobile laptop with iwl5300 NICs is first fixed at one location and then moved to another. We repeated this process and placed the device at 10 different positions respectively. Fig. 11(a) illustrates the cumulative distribution (CDF) of localization errors across the 10 positions. In our experiments, for over 90% of data points, the localization error falls within the range of 1 meter, and the 50% accuracy is less than 0.5 m. In such a dynamic environment with lots of factors interfering the propagation of signals, FILA exhibits a preferable property indicating that the fine-grained nature of CSI is beneficial to improve the accuracy of corresponding RF-based approach.

Fig. 11(b) depicts the cumulative distribution function errors per location detected at the university lecture theatre. Even for such a much larger space, FILA can locate objects in the range within 1.8 meters of their actual position with 90 percent probability, which is acceptable for most location-based applications.

Across the above two typical single room indoor scenarios, FILA achieves median accuracy of 0.45 m and 1.2 m, respectively. It is therefore safe to conclude that the proposed CSI-based scheme performs much better than RSSI-based one when locating objects in a typical indoor building with multipath effect.

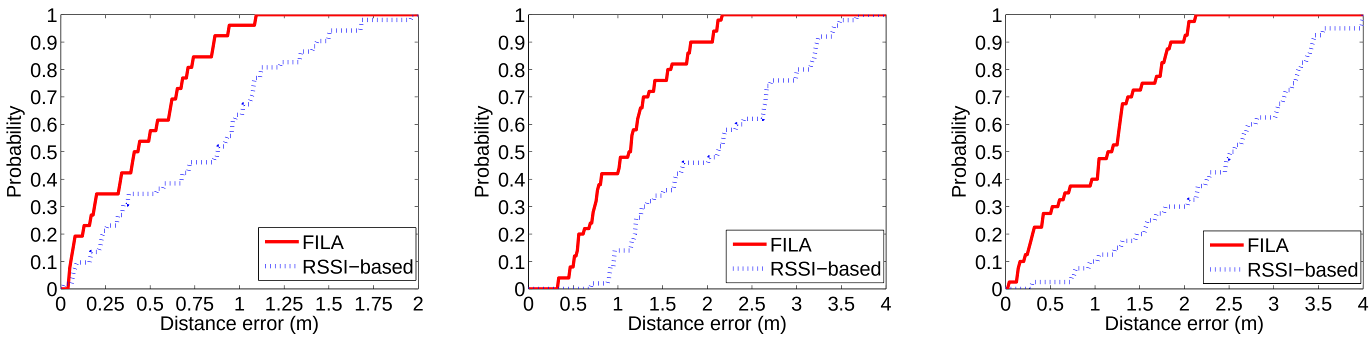
Localization Accuracy in Multiple Rooms

In our previous experiments, the APs and client are placed in the same room. We also examined the corridor scenario where several APs are deployed in the multiple rooms. Specially, we take into consideration both the complicated multipath effect and the shadowing fading brought by wall shield. We first fix the position of the object at some reference nodes, and collects the AP coordinate and CSI value for offline training. Then, we move the object to arbitrary positions for online tracking. The moving speed is around $1m/s$ and we collect 20 CSIs and RSSIs at each position.

In Fig. 11(c), we plot the cumulative distribution of location errors across 10 positions. It is shown that multipath propagation does degrade the accuracy of object localization as well as the shadowing in the multiple rooms scenarios. However, FILA is robust enough to maintain the degradation. More importantly, FILA can achieve median accuracy of 1.2 m in this corridor environment. This result indicates that FILA is able to effectively estimate and compensate for gain differences across multiple rooms.

6.2.2 Latency of CSI-based propagation model localization

Two main phases contribute to the latency of FILA, the calibration phase and location determination phase. Since the environment factor and the fading exponent vary in different environments, we need to conduct calibration to train these two parameters for the refined propagation model. We should collect CSIs at some pre-known positions to calculate these two parameters using our fast training algorithm. Actually, this process can be finished before localization as an offline task since the APs can use each other's information for the calibration. In our FILA system, the AP takes about $0.8ms$ to transmit a packet with 100 bytes beacon message in



(a) Research Laboratory (b) Lecture Theatre (c) Corridor

Fig. 11: CDF of localization error in different indoor environments.

IEEE 802.11n. Each time we collect 20 CSIs and the time will be $0.8 \times 20 = 16ms$. The calibration process can be done within $2ms$ according to our measurement on a HP laptop with 2.4GHz dual-core CPU. In the location determination phase, the IFFT and FFT process can leverage the according hardware blocks in the wireless NICs whose running time is ignorable. While we conduct these signal processing on laptop consuming around $2ms$, including the time needed for the trilateration location calculation. Therefore, the time consumption for both training and location determination is within several ms. In our experiment, the walking speed of user is around $1m/s$. The average tracking error is around $1.2m$ as shown in Figure 11(c). For multiple rooms environment, the simple alpha-tracking algorithm [12] can be applied for triggering the training. We only need to train these parameters once unless the environment changes greatly. In summary, our system can reach the average time tracking latency to as fast as about $0.01s$, which significantly outperforms previous RSSI-based tracking systems [4] (usually $2-3s$).

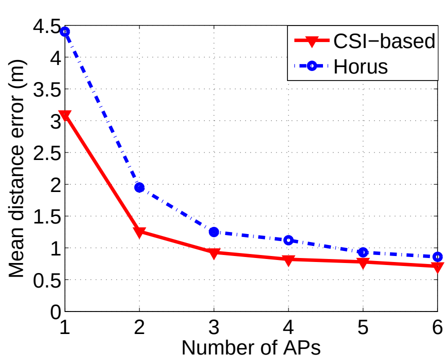


Fig. 12: Mean distance error.

6.3 Performance Evaluation of CSI-based Fingerprinting

6.3.1 Accuracy

First, we evaluate the accuracy of the proposed CSI-based probability algorithm and compared with Horus, the widely used RSSI-based fingerprinting system. The mean distance error in the two different environments is shown in Fig. 12. Our evaluation shows that CSI-based method can achieve the median accuracy of $0.65m$, which outperforms Horus by about $0.2m$, and the gain is about 24%. Moreover, in the corridor scenario, where covered by 6 APs and 3 APs were taken into computation, the mean accuracy of our approach is $1.07m$ which is $0.35m$ lower than Horus system, about 25% gain over Horus.

In addition, we compare the two approaches concerning different numbers of APs in corridor. Fig. 14 depicts the average accuracy according to the amount of APs varying from 1 to 6. Since richer information to estimate the location can be obtained from the more APs, both lines demonstrate the accuracy improvement. In particular, our approach reduced the mean distance error by 29% on the average. Obviously, these results show the effectiveness of the proposed CSI-based location system and indicate the benefits from indoor environment with dense-deployed APs. When the AP is sparse, our scheme performs much better than the RSS-based one.

6.3.2 Precision

Fig. 15 illustrates the cumulative distribution (CDF) of localization errors in the laboratory. The data were collected across the 28 positions in the laboratory. From Fig. 15, we can observe that the confidence interval with confidence level 90% for the error is (-1.3m, +1.3m), which means the CSI-based localization error falls within the range of 1.3 meters, and the 50% accuracy is less than 0.6 m. However, the Horus can locate objects in the range within 1.6 meters of their actual position with 90 percent probability, and the median accuracy is 0.8.

Unlike the first scenario that 3 APs and client are placed in the same room, we also examined the corridor testbed where the 6 APs are deployed in the multiple rooms. Fig. 13 depicts the cumulative distribution of positioning errors across 20 positions. We can easily observe that both our approach and Horus can achieve the median accuracy less than $1.25m$. However, the accuracy improvement of our approach over Horus for 90% of

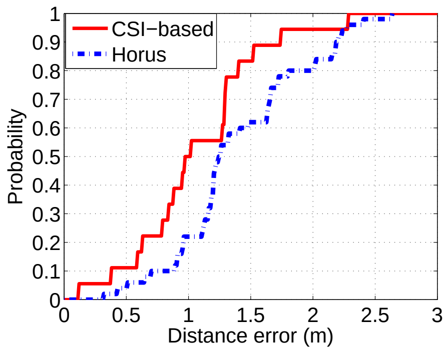


Fig. 13: CDF of localization error in Corridor.

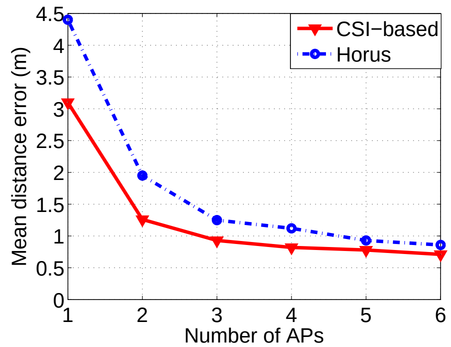


Fig. 14: Mean distance error with different numbers of APs.

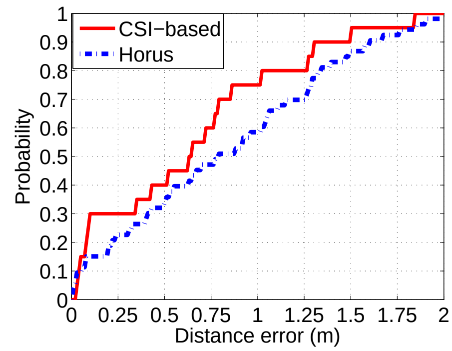


Fig. 15: CDF of localization error in Laboratory.

data points is $0.55m$. We can conclude that our approach exhibits a preferable property since the fine-grained and frequency diversity nature of CSI is beneficial to improve the precision of location fingerprinting.

7 Conclusions and Future Work

Localization is one of the most appealing applications and becomes increasingly common in our daily life. RSSI-based schemes have been widely used to provide location-aware services in WLAN. However, in this paper, we observe that RSSI is roughly measured and easily affected by the multipath effect which is unreliable. We then use the fine-grained information, that is, Channel State Information (CSI), which explores the frequency diversity characteristic in OFDM systems to build the indoor localization system FILA. In FILA, we process the CSI of multiple subcarriers in a single packet as effective CSI value CSI_{eff}, and develop a refined indoor radio propagation model to represent the relationship between CSI_{eff} and distance. Based on the CSI_{eff}, we then design a new fingerprinting method which leverages the frequency diversity. To demonstrate the effectiveness of FILA, we implemented it on the commercial 802.11n NICs. We then conducted extensive experiments in typical indoor environments and the experimental results show that the accuracy and speed of distance calculation can be significantly enhanced by using CSI.

In this work, we just use the simplest trilateration method to illustrate the effectiveness of CSI in indoor localization. The future research in the new and largely open areas of wireless technologies can be carried out along the following directions. First, we can leverage the available multiple APs to improve the location accuracy in some extent. Second, in this paper we only leverage the frequency diversity, however, the spatial diversity can also be exploited in MIMO to enhance the indoor localization performance. Third, since some of the smart phones have 802.11n chipset, the next step of our work is to implement FILA in smart phone.

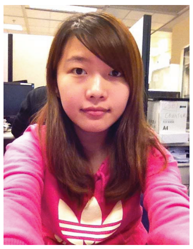
Acknowledgments

This research was supported in part by the 2012 Guangzhou Pearl River New Star Technology Training Project, Hong Kong RGC Grants HKUST617811, 617212, S&T Project of Guangdong Province, China Grant No.2011A011302001, NSFC Grant No. Grant No. (60933011 and 61027009), NSFC-Guangdong Joint Fund (U0835004), the National Key Technology R&D Program (2011BAH27B01, 2011BHA16B08), the Project Science and Technology of Guangdong Province (2011168014 and 2011912004) and the Major Science and Technology Projects of Guangdong (2011A080401007).

References

[1] P. Bahl and V. N. Padmanabhan, “Radar: an in-building rf-based user location and tracking system,” in *Proc. of IEEE INFOCOM*, 2000.
[2] J. Liu, Y. Zhang, and F. Zhao, “Robust distributed node localization with error management,” in *Proc. of ACM MobiHoc*, 2006.
[3] D. Moore, J. Leonard, D. Rus, and S. Teller, “Robust distributed network localization with noisy range measurements,” in *Proc. of ACM Sensys*, 2004.
[4] D. Zhang, J. Ma, Q. B. Chen, and L. M. Ni, “An rf-based system for tracking transceiver-free objects,” in *Proc. of IEEE PerCom*, 2007.
[5] D. Zhang and L. M. Ni, “Dynamic clustering for tracking multiple transceiver-free objects,” in *Proc. of IEEE PerCom*, 2009.
[6] D. Halperin, W. J. Hu, A. Sheth, and D. Wetherall, “Predictable 802.11 packet delivery from wireless channel measurements,” in *Proc. of ACM SIGCOMM*, 2010.
[7] A. Goldsmith, *Wireless Communications and Networks:3G and beyond*. Tata McGraw Hill Education Private Limited, 2010.
[8] K. Fazal and S. Kaiser, *Multi-carrier and spread spectrum systems : from OFDM and MC-CDMA to LTE and WiMAX*. Wiley, 2008.
[9] M. Youssef and A. Agrawala, “The horus wlan location determination system,” in *Proc. of ACM MobiSys*, pp. 205–218, 2005.
[10] S. Fang, T. Lin and K. Lee, “A novel algorithm for multipath fingerprinting in indoor wlan environments,” *IEEE Trans. Wireless Commun.*, 2008.
[11] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, “Practical robust localization over large-scale 802.11 wireless networks,” in *Proc. of ACM MobiCom*, pp. 70–84, 2004.
[12] T. He, S. Krishnamurthy, T. Luo, L. andYan, K. B., L. Gu, R. Stoleru, G. Zhou, Q. Cao, P. Vicaire, J. A. Stankovic, T. F. Abdelzaher, and J. Hui, “VigilNet: An Integrated Sensor Network System for Energy-Efficient Surveillance.,” in *ACM Transactions on Sensor Networks*, 2006.

Kaishun Wu is currently a research assistant professor in Fok Ying Tung Graduate School with the Hong Kong University of Science and Technology. He received the Ph.D. degree in computer science and engineering from Hong Kong Uinversity of Science and Technology in 2011. His research interests include wireless communication, mobile computing, wireless sensor networks and data center networks.



Jiang Xiao is currently a first year Ph.D student in Hong Kong University of Science and Technology. Her research interests mainly focused on wireless indoor localization systems, wireless sensor networks, and data center networks.

Youwen Yi received his B.Sc (Hons) degree from Harbin Institute of Technology in 2007, and receive his M.Sc degree from Huazhong University of Science and Technology in 2009. From 2009 and 2011, he was a research assistant at Hong Kong University of Science and Technology . Currently, he is a senior research engineer at Huawei Technologies. His research interests include real-time indoor localization system, and next generation self-organized network.

Dihu Chen received B. Sc. and M. Phil. degree from the Semiconductor Physics of Sichuan University in 1986 and 1989, respectively, and he received Ph.D. degree in solid state electron from Department of Electronic Engineering, The Chinese University of Hong Kong in Dec. 2000. Now he is a professor, head of the department and ASIC Design Center of Sun Yat-sen University. He is currently working on electronic devices, IC design and design methodology.

Xiaonan Luo is a Professor in the School of Information Science and Technology, Director of the Computer Application Institute, Sun Yat-sen University, China. His research interests include mobile computing, computer graphics and CAD.

Lionel M. Ni is Chair Professor in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology (HKUST). He also serves as the Special Assistant to the President of HKUST, Dean of HKUST Fok Ying Tung Graduate School and Visiting Chair Professor of Shanghai Key Lab of Scalable Computing and Systems at Shanghai Jiao Tong University. A fellow of IEEE, Dr. Ni has chaired over 30 professional conferences.