IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 1, 2017, accepted December 9, 2017, date of publication December 27, 2017,
date of current version March 12, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2785789

An Efficient Construction of Confidence Regions via Swarm Intelligence and Its Application in Target Localization

FRANK PO-CHEN LIN AND FREDERICK KIN HING PHOA
[1]College of Electrical Engineering and Computer Science, National Taiwan University, Taipei 106, Taiwan
[2]Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan

Corresponding author: Frederick Kin Hing Phoa (fredphoa@stat.sinica.edu.tw).

This work was supported in part by the Career Development Award of Academia Sinica, Taiwan, under Grant 103-CDA-M04 and in part by the Ministry of Science and Technology, Taiwan, under Grant 105-2118-M-001-007-MY2.

ABSTRACT It is essential to enhance the speed and accuracy of the localization process to gain the robustness and instantaneous properties and to adapt from the practical environment of a confidence band. In this paper, we proposed a new received signal strength indicator-based method to construct a real-time confidence band, which was composed by multiple confidence region sets in a multivariate normal distribution, associated to a target's trajectory for location-based services. Based on the concept of weighted positioning circular algorithm, we designed a new objective function to take into consideration the signal disruptions of the surrounding environments. The characteristics of the state of motion for the moving target were then inferred from the status of each confidence region. In order to speed up the localization process to obtain the real-time estimate of the confidence band via our objective function, we proposed in this paper a swarm intelligence-based localization optimization algorithm, which was modified from the standard framework of a novel swarm intelligence-based evolutionary algorithm.

INDEX TERMS Confidence band, localization, swarm intelligence based (SIB) optimization, wireless sensor network (WSN), received signal strength indicator (RSSI), circular positioning algorithm.

I. INTRODUCTION

Wireless communication has been widely used in our daily life after some remarkable recent breakthroughs [1]. The widespread deployment of wireless devices allows the researchers to study the feasibility of utilizing Radio Frequency (RF), like the Wi-Fi being standardized by IEEE 802.11, to provide Location Based Services (LBS) as well as communication services to users. As the most well-known RF, the Wi-Fi uses a specific frequency range within the signal bands of both 2.4 GHz and 5 GHz as a channel and serves as a medium to allow devices to connect to the Wireless Local Area Network (WLAN), which provides wireless network communication within a limited area.

Rather than passively listening to the signals, the Wi-Fi-based devices mostly apply the technique of active scanning to gain access to the Internet through WLAN. In specific, these devices scan through each channel, send a Probe Request (PR) management frame and ask for the available network on that channel to the nearly Wi-Fi access points (APs). Specific characteristics, including Received Signal Strength Indicator (RSSI), Time of Arrival (TOA) and Time Difference of Arrival (TDOA), can be extracted from each PR. Consequently, these characteristics are used to obtain an estimate of the device position via localization algorithms [2], but it is not trivial to model the radio propagation in various environments due to diffraction, scattering, shading, severe multi-path, low probability for availability of line-of-sight (LOS) path and specific site parameters such as moving objects and numerous reflecting surfaces [3].

Many localization algorithms, some are range-based and others are range-free [4], are proposed in literatures. In general, a range-based algorithm provides more accurate localization results than a range-free algorithm, so it is traditionally suggested for estimating positions, but it is not trivial to implement due to high hardware requirements. In specific, the TOA-based localization algorithms require precise clock synchronization [5]–[9], the TDOA-based localization algorithms require ultrasonic transceivers [10]–[14] and

the AOA-based localization algorithms require directional antenna [15]–[17]. On the contrary, no additional hardware is required for RSSI-based localization algorithms, thus it has been broadly used in many applications for its simplicity and low implementation cost [4]. Most radio propagation models in RSSI-based localization techniques assume to have perfect channels, but it does not hold in practice. [18] proposed a RSSI-based weighted circular algorithm to take care of imperfect channels, but it has no guarantee on the real-time position estimate of a target. In order to complete the localization process within a limited amount of time, new methods with speed and accuracy optimization are of great interest.

Swarm Intelligence Based (SIB) method, which is famed for its fast convergence, simple implementation and low computational cost, comes into play. It was first proposed to solve the optimization problems in the designs of experiments in statistics [20]–[22]. No assumption is required on the objective function of SIB. Due to the architecture in the SIB algorithm, its computational performance can be doubled by applying CPU parallel computing on a single dual-core computer [19]. On the other hand, some have explored device trajectory along different time-slots to gain a better resolution into the target device. Trajectory data is considered as a core revealing details of instantaneous behaviors piloted by mobile entities [23]. Although most of the researches in localization focus on the techniques for positioning, the characteristics of the trajectory formed by user's position through different time slots is seldom mentioned and discussed. These characteristics from feature extraction of the device's estimated position are important for providing promising LBS.

The confidence band is well-known in statistics to represent the uncertainty in a function estimation from a limited or noisy data [24]–[26], [30], [31]. Various informative analyses and different dependency relationships among factors can be extracted by the construction of the confidence band. Such an important tool has not been popularly applied in the field of engineering yet. This leads to our proposed method to construct real-time confidence bands for a target's trajectory based on a new RSSI localization approach. In specific, we introduce a new objective function based on the concept of the weighted positioning circular algorithm to consider the conditions in the real environment. We then modify the SIB method to search for a good solution via speeding up the process of localization for a duplicate sampling of distance estimation. A confidence band for the target's moving trajectory is constructed via simultaneous inference and coordinate dependency of the estimated target position. Finally, the characteristics of the state of motion are obtained from the status of the confidence band.

The paper is organized as follows. In section II, the background of channel modeling localization and the concept of confidence region are described. Section III presents the approach of the proposed localization algorithm and the construction of trajectory-based confidence band. Section IV evaluates the result of our proposal. Section V concludes the paper and outlines the future work.

II. CHANNEL MODEL LOCALIZATION AND CONFIDENCE REGION

A. CHANNEL MODELING LOCALIZATION

Consider two types of nodes in a Wi-Fi network: mobile and APs. As the mobile tries to gain access to the internet, our device automatically sends a PR management frame searching for the available network on that channel actively to nearby APs. On the other hand, the nearby APs obtains the RSSI of the requests sent out by our mobile devices. The distance between the device and the APs can be calculated by the APs by applying various channel models according to the need of applications and assumptions [27]. This paper considers the Free-space path loss (FSPL) model: $\bar{PL}(d_i) \propto (d_i/d_o)^{\eta}$ and $\bar{PL}(dB) = \bar{PL}(d_o) - 10\eta\log_{10}(d_i/d_o)$ for $i = 1, 2, \ldots, n$, where d_o is the near earth reference distance, $\bar{PL}(d_o)$ is the signal strength at distance d_o and η is the signal attenuation factor with value between 2 to 6 according to different environments. It is the most common model [4] to obtain the corresponding distance between the device and the APs. The position of the device is then determined from the estimated distances by using a multi-lateration algorithm. The second equation represents the channel model with perfect channel condition without considering the attenuation caused by flat fading effect. In this paper, the case of shadow fading or slow fading is considered. The noise is considered as a random variable from a Gaussian distribution with a standard deviation of σ. Then the RSSI-based weighted circular positioning algorithm [18] is applied to estimate the position of the device in this paper.

In a circular positioning algorithm [28], the target's location is determined from the intersections of circumferences centered at the APs. The positions of the APs are known and located accurately in a certain position, and the radius of the circumferences is estimated by the relationship between RSSI in the Free-space path loss model. Due to the existence of uncertainty in the RSSI-distance estimation caused by the channel environment, the intersections of the circumferences is presented as a range of area instead of a single unique point. A nonlinear least square method is applied to solve the problem as it provides a comparably more accurate result than other methods [29]. Our goal is to minimize the sum-square error on the device location, thus our objective function is expressed as $\varepsilon = \sum_{i=1}^{N} (\sqrt{(x_i - x)^2 + (y_i - y)^2} - \hat{d}_i)^2$ where x_i and y_i are the x- and y-coordinates of the ith AP, and $\hat{d}_i$ is the estimated distance from the target to the ith AP.

In real localization applications, the channel environment may vary under different circumstances and cause a considerable amount of variation with respect to the standard deviation of the Gaussian noise, generating from shadow fading and model error caused by uncertain propagation parameters. These errors in the RSSI-distance conversion may lead to an inaccurate representation of the real radio channel and produce undesirable results in localization. The second equation of FSPL model reveals that the variance of the Gaussian noise is correlated with the distance between the target device and

the AP. Since a shorter distance between the target device and the AP results in a higher accuracy of the estimated distance, a majorization approach on the accurate measurement is able to obtain an accurate localization result.

In light of above, a weighted circular positioning algorithm [18] is a good substitute to its standard version in this manner. The error in the distance estimation for the ith AP is given by $e_i = d_i - \hat{d}_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} - \hat{d}_i$ and the weighted least square error is $\varepsilon = e^T S^{-1} e$, where S is the covariance matrix of the random vector. Since the elements in the random vector is assumed to be uncorrelated, so S can be simplified to a diagonal matrix. The objective function can then be modified as $\varepsilon = \sum_{i=1}^{N} \frac{e_i^2}{Var(e_i)}$. The variance of the estimated distance $\hat{d}_i$ can be expressed as $Var(\hat{d}_i) = \exp(2\mu_d)(\exp(2\sigma_d^2) - \exp(\sigma_d^2))$, where $\mu_d = \ln d_i$ and $\sigma_d = \frac{\sigma \ln 10}{10\eta}$.

Finally, the objective function can be expressed as

$$f(x) = \sum_{i=1}^{N} \frac{1}{Var(\hat{d}_i)} \left(\sqrt{(x_i - x)^2 + (y_i - y)^2} - \hat{d}_i\right)^2. \quad (1)$$

As shown in Equation 1, the objective function is equal to the weighted sum of the squared errors in the set of estimated distances and the weights are given according to the variation of $\hat{d}_i$. Since an estimate of a large distance results in a large variance, the objective function with ill-assigned weight would lead to unreliable influence on the measurements.

B. CONFIDENCE REGION

Confidence intervals were introduced for estimating a single parameter for random variables of single-dimension for which the difference between the parameter and the observed estimate is not statistically significant at a specific level. Confidence regions can be similarly defined for vectors of parameters, such as the mean vector μ for multivariate random vectors. The resulting confidence regions are in the form of ellipsoids. The regions cover the vector μ with a specific level of assigned probability, usually high. To efficiently obtain the confidence region of a mean vector, the correlations between the estimate parameters are taken into consideration. According to the Central Limit Theorem, the joint probability distribution of the estimated locations is reasonably assumed to be normally distributed. Consequently, construction of a confidence region for a multivariate normal distributed random vector is recalled as follows.

Consider n i.i.d p-dimensional random vectors $X_i = \{x_{i1}, x_{i2}, \ldots, x_{ip}\} \sim N_p(\mu, \Sigma)$ for $i = 1, 2, \ldots, n$ and μ is the p-dimensional vector $\mu = E[X_i] = [E[x_{i1}], E[x_{i2}], \ldots, E[x_{ip}]]^T$. For $x \in \mathbb{R}^p$, the joint probability density function (pdf) of X_i is $f(x|\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{1}{2}p} |\Sigma|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\}$. It is necessary to derive the mean estimate $\bar{X}$ and the covariance matrix S in order to construct a confidence region from the estimated locations, and they can be obtained by using the maximum likelihood estimation (MLE) of their pdf. Given the observations of a k-dimensional random vector $x^k = \{x_1, x_2, \ldots, x_k\}$, where $x_k \sim k_N(\mu, \Sigma)$, the estimate $\bar{x}$ can be expressed as $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ and the estimate $\hat{\Sigma}' = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T$. Note that $\hat{\Sigma}'$ is biased estimator of Σ, the unbiased estimator of the sample covariance matrix is $\hat{\Sigma} = \frac{N}{N-1} \hat{\Sigma}'$. Since $\bar{x}$ and $\hat{\Sigma}$ are independent, the former follows the normal distribution $\sqrt{N}(\bar{x} - \mu) \sim N_p(0, \Sigma)$ and the latter follows the Wishart distribution $(N - 1)\hat{\Sigma} \sim W_p(N - 1, \Sigma)$. It follows that

$$\frac{N - p}{(N - 1)p} N(\bar{x} - \mu)^T \hat{\Sigma}^{-1} (\bar{x} - \mu) \sim F_{p,N-p} \quad (2)$$

where $F_{p,N-p}$ is the F-statistics of degrees p and $N - p$. A confidence region $A \in \mathbb{R}^p$ corresponds to an area with coverage probability of $1 - \alpha$ that contains the mean vector μ, where $\alpha \in (0, 1)$ if $P(\mu \in A) \geq 1 - \alpha, \forall \mu \in \mathbb{R}^p$. Therefore, the $(1 - \alpha)100\%$ confidence region for the mean of a p-dimensional normal distribution is

$$P(N(\bar{x} - \mu)^T \hat{\Sigma}^{-1} (\bar{x} - \mu) \leq \frac{(N - 1)p}{N - p} F_{p,N-p}(\alpha)) = 1 - \alpha \quad (3)$$

and it corresponds to the set $A = \{\mu \in \mathbb{R}^p : N(\bar{x} - \mu)^T \hat{\Sigma}^{-1} (\bar{x} - \mu) \leq \frac{(N-1)p}{N-p} F_{p,N-p}(\alpha)\}$. The inequality defines a p-dimensional ellipsoid centered at $\bar{x}$.

III. CONSTRUCTION OF CONFIDENCE BAND

This section introduces the modified objective function for duplicate RSSI sampling and the SIB evolutionary algorithm, followed by the construction of confidence band of the target's moving trajectory.

A. OBJECTIVE FUNCTION FOR DUPLICATE SAMPLING

To improve the accuracy of localization, we consider the average of the RSSI signals sent from the targets to estimate their distances between the APs. The estimated distance between a target and each AP is estimated in multiple times. Each acquired distance is considered as a sample of the estimation. The amount of sampling for the estimated distance between the target and each AP varies according to the state of motion of the target. If the target moves slowly or stays in the same location, the number of samples for each estimation is expected to be large, and vice versa. After the estimations, the mean of the estimated distance can be obtained by iteratively sampling the distances between the target and each AP as $\bar{\hat{d}}_i = \frac{\sum_{i=1}^{n_i} \hat{d}_i}{n_i}$, where $\hat{d}_i$ is the estimated distance between the target and the ith AP, n_i is the number of estimated samples between each AP and $\bar{\hat{d}}_i$ is the mean value for $\hat{d}_i$.

$\bar{\hat{d}}_i$ is applied to the localization objective function as the final estimated distance between the target and the ith AP. The intention for this approach is to reduce the variance of each estimation and obtain an accurate result according to the law of large numbers. However, a huge amount of distance estimations is required in order to obtain the variance of $\bar{\hat{d}}_i$ if the objective function for the weighted circular algorithm

is directly applied. This leads to a large sampling time for estimating the distances $\hat{d}_i$ in order to get enough amount of $\hat{\bar{d}}_i$ to obtain the corresponding variance $Var(\hat{\bar{d}}_i)$. It turns out to be time-consuming and inefficient for a real-time localization for a moving target, which makes the construction of confidence band unrealistic. Consequently, this paper modifies the objective function of the algorithm by directly approximating the variance of the sample mean $Var(\hat{\bar{d}}_i) = \frac{Var(\hat{d}_i)}{n_i}$.

On the other hand, the variance of $\hat{d}_i$ in the weighted circular algorithm was assumed to follow a logarithmic normal distribution under mathematical assumptions of the channel model. However, this may result in some distortion and inaccurate estimates of the real variances. As the estimation is different under different environments, it is appropriate to find the variances based on the real conditions instead of considering to assume a general distribution. In other words, the variance of $\hat{d}_i$ can be obtained from the sampled values. The unbiased estimator of the variance of the sample distances is $Var'(\hat{d}_i) = S_i = \sqrt{\frac{1}{n_i-1}\sum_{j=1}^{n_i}(\hat{d}_i - \bar{\hat{d}}_i)^2}$, and the variance of the sample mean is $Var'(\hat{\bar{d}}_i) = \frac{S_i}{n_i}$. Thus, the final objective function used in our method is expressed as

$$f(x) = \sum_{i=1}^{N} \frac{n_i}{Var'(\hat{d}_i)} \left(\sqrt{(x_i - x)^2 + (y_i - y)^2} - \bar{\hat{d}}_i\right)^2. \quad (4)$$

In most conditions, n_i would be the same and the equation can be simplified. Nevertheless, the stability of signal detection due to signal attenuation for Wi-Fi APs located with a certain amount of distance beyond the target should be taken into consideration. While most of the existing Wi-Fi based localization approaches aim to find a good estimate of the target's location, few have taken the stability of faraway Wi-Fi APs into consideration. This phenomenon becomes statistically significant when the target is moving under a particular velocity and it leads to inaccurate results when the band is constructed. The proposed objective function is capable of handling this problem by slightly adjusting the weights to gain an extra accurate estimate.

The modification of our objective function in Equation 4 have three major contributions: First, the variance of the sample mean $Var(\hat{\bar{d}}_i)$ can be obtained approximately without requiring a large amount of estimated samples of distances. Second, the assigned weights to different estimates are more accurate and closer to the actual condition in practice. Third, the objective function has taken the unstable conditions of signal attenuation into consideration.

B. THE SIB ALGORITHM FOR LOCALIZATION

The SIB algorithm was first proposed to optimize designs of experiments in statistics under various criteria of design properties [20]. This method can be applied for the search of both continuous and discrete domains. The general framework of the SIB algorithm is shown in Figure 1.

In the step of initialization, possible solutions are generated as initial particles, the values of the objective function for these particles are evaluated, each of which has its own perceived location of initial optimum in the search space called Local Best (LB) particles. All particles share information by comparing its initial LB with other to perceive the overall optimum location called Global Best (GB) particle. In this step, the number of the initial particles and the stopping criterion are set. For the particles to collectively arrive at the perceived overall optimum location, they go through the steps of MIX and MOVE operations iteratively after the initialization step. LB particles and the GB particle are updated continuously in each iteration until the stopping criterion is fulfilled, which can be the reach of either the pre-specified maximum number of iterations or a known optimal value of the GB particle.

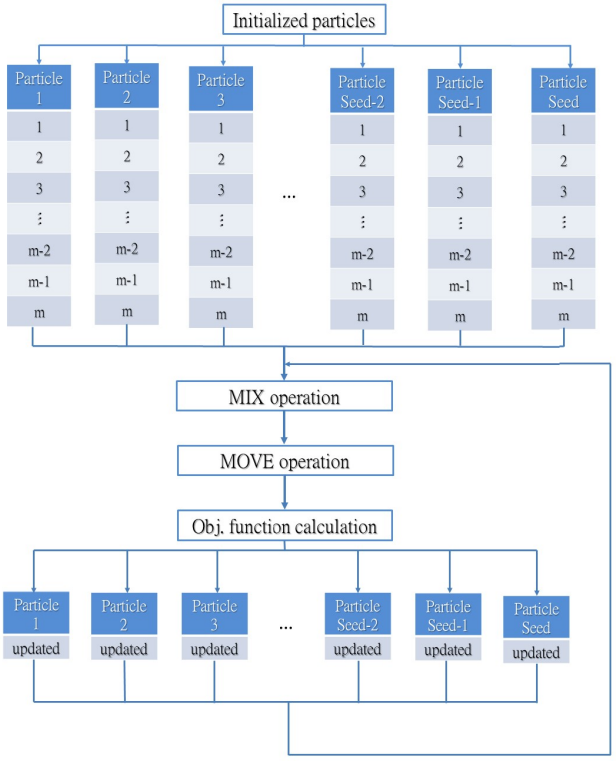

FIGURE 1. The process of SIBL algorithm.

In the MIX operation, each particle goes through the deletion and addition steps. Assume that every particle X is consisted of m discrete units: $X = \{x_1, x_2, \ldots, x_m\}$. The deletion step for X is a recursive procedure that determines the best units to be removed from X. It continues iteratively with a single element being removed in each iteration. The number of deleted units is pre-specified in the initialization step. The reduced particle is denoted by X_{-R}. We then consider the good particle Y chosen from the LBs or the GB. The addition step iteratively adds a single element from Y into the reduced particle X_{-R} in a similar fashion. When the MIX operation of X is completed, two new candidate particles are resulted: *mixwGB* when Y is the *GB* particle, and *mixwLB* when Y is the *LB* particle.

The MOVE operation is a decision-making procedure to select the best particles among all candidates with the optimal values of objective functions. Three candidates among X, *mixwLB* and *mixwGB* are compared and chosen for update. If both *mixwLB* and *mixwGB* have worse objective values than X, some units of X are randomly chosen and replaced by some other random units to avoid local optimal solution.

Now we consider the SIB method for localization problems, namely the SIBL algorithm. We considers a square domain for localization with the center of the area being the origin (0, 0). In the step of initialization, the size of the square domain $\ell \times \ell(m^2)$ is first determined. Then initial particles are generated and designed in the form of two-dimensional vectors $x_k = \{x_{k_1}, x_{k_2}$ for $k = 1, 2, \ldots, m$. Each element is assigned with a random value within the range of $[-\frac{\ell}{2}, \frac{\ell}{2}]$ to represent the target's location in the Cartesian coordinates. Since the objective values for the new objective function has yet been defined theoretically, this paper only considers the number of maximum iterations as the stopping criterion. Based on our observations and experience, the objective value converges rapidly with an average of three iterations in the experiment, so we assign the maximum number of iterations to be five, which is slightly higher to avoid incomplete convergence. On the other hand, the number of elements in each designed particle is small comparing to that in the SIBSSD algorithm [22], so the execution speed is fast as the number of discrete units to exchange in the MIX operation is limited to either one or two. Since the execution time for the MIX operation dominates the running time of the entire algorithm, fast speed in MIX provides the capability for the SIBL algorithm to localize the target rapidly and supports the need for real-time positioning to construct the confidence band.

However, there are some tradeoffs behind the benefits brought by the design of particles. We denote the number of discrete units to be exchanged in the MIX operation as q. If q is fixed to be one, only one element is being exchanged in each iteration of the entire MIX operation, which limits the diversity and improvement of the particles. In contrast, if q is assigned with a value of two, all elements in each particle are exchanged. The potentially ''good'' elements in the original particle can hardly be maintained. In order to increase the diversity of particles for localization while retaining the speed performance for the SIBL algorithm, the MIX operation for SIBL is slightly modified so that it does not change the design of particles. For the MIX operation in SIBL, the value of q is determined dynamically instead of being directly assigned with a fixed number. The value of q is assigned to either one of its possibilities with each having an equal probability of 0.5. Throughout this process, the diversity of the generated particles during the MIX operation can be increased whereas the original design of the particle still remains the capability to construct the real-time confidence band for the target's trajectory. In the SIBL algorithm, the MOVE operation directly inherits the concept introduced in the standard framework of the SIB algorithm.

The pseudo code of the SIBL algorithm is as shown below.

1: Determine the size of the square domain $\ell \times \ell(m^2)$.
2: Randomly generate a set of initial particles.
3: Evaluate the localization-based objective function value for each particle.
4: Initialize the LB for all particles.
5: Initialize the GB.
6: **while** not converge **do**
7: Determine the number of elements to be exchanged.
8: For each particle, perform the MIX operation.
9: For each particle, perform the MOVE operation.
10: Evaluate the localization-based objective function value for each particle.
11: Update the LB for all particles.
12: Update the GB.
13: **end while**

C. CONSTRUCTION OF THE CONFIDENCE BAND

Confidence band can be categorized into many different types, each type is suitable for a specific usage or an application. In this paper, the confidence band corresponds to the simultaneous confidence band. The simultaneous confidence band is formed by a collection of confidence sets. With a given coverage probability, the simultaneous confidence band covers the corresponding true values for all collection of confidence sets. To construct the confidence band of moving trajectory, the confidence sets of every estimated position in each area is established in advance. The confidence sets for every estimated area are considered as confidence regions of a multivariate normal distribution of two dimensions. A confidence region of a certain area is then constructed when the target's position is detected within. To construct the confidence regions, we consider N estimates of positions for a certain area with a total number of areas M. For $i = 1, 2, \ldots, M$, each estimated position of a particular area, namely the ith area, is considered as a two-dimensional normal distributed random vector $x_{ij} = \{x_{ij_1}, x_{ij_2}\}$ for $j = 1, 2, \ldots, N$. The mean vector and the covariance matrix can be obtained by $\bar{x}_i = \frac{1}{N}\sum_{j=1}^{N} x_{ij}$ and $\hat{\Sigma}_i' = \frac{1}{N}\sum_{i=1}^{N}(x_{ij} - \bar{x})(x_{ij} - \bar{x})^T$ respectively. Then the confidence region of the estimated positions of a certain location can be obtained from

$$P(N(\bar{x}_i - \mu)^T \hat{\Sigma}^{-1}(\bar{x}_i - \mu) \leq \frac{(n_i - 1)p}{n_i - p} F_{p,n_i-p}(\alpha)) = 1 - \alpha \tag{5}$$

where α is the factor to determine the confidence level for each region $A = \{\mu \in \mathbb{R}^2 : n_i(\bar{x}_i - \mu)^T \hat{\Sigma}^{-1}(\bar{x}_i - \mu) \leq \frac{(n_i-1)p}{n_i-p} F_{p,n_i-p}(\alpha)\}$.

We construct the confidence band after the confidence regions for all domain areas are established and the targets are appeared near the APs. In specific, Equation 5 is satisfied for all $x_i \in X$ the set of positions in all areas of localization, $i = 1, 2, \ldots, M$, so that the constructed confidence band covers the true values for all collection of confidence sets. In other words, the confidence band can be recast as the joint confidence regions constructed by adjusting the confidence

levels of every confidence region. In this paper, we apply the Bonferroni method to adjust the M confidence regions such that the confidence level of each of the M regions is increased from $1-\alpha$ to $1-\alpha/M$. As a result, the probability that the simultaneous confidence band covers the corresponding true values for all collection of confidence sets is approximately $1-\alpha$ and Equation 5 is satisfied.

The confidence band is informative if the number of confidence regions are large. However, since there are inappropriate positions for the deployment of APs, Wi-Fi based localization cannot always target all positions of every area along the entire traveling path. Since we construct the confidence band based on the target's trajectory, the estimation accuracy of every position need not always be precise and a certain amount of tolerance is allowed. In order to construct a well-informed confidence band that covers as many confidence regions as possible, the position estimate for the areas with less than three intersections of AP signals is targeted by the APs to build up the band. The established confidence region under this situation is expected to have a large size originated from the increase of estimation variation. Such minor defects are acceptable to the whole structure of the confidence band.

Sometimes the client sends out a PR at its lowest supported data rate, typically 1 Mbps, the rate of receiving RSSI signals as well as the number of estimations for the distance between AP and target are assumed to be constant. Then the confidence band for a traveling trajectory reveals the average speed of the target along the path without any other extra hardware implementations. Since the APs localize the moving target repeatedly, the moving speed depends on the number of localization operated by the APs for each position. The result show that the size of the confidence band is highly correlated with the average speed of the moving target. Subtle changes in the moving speed results in a non-neglectable alteration of the size of the confidence band. This fact implies that the confidence band provides important information about the motion of the traveling target, which can be useful on many applications after appropriate data analysis is conducted.

IV. SIMULATION AND ANALYSIS

The simulation performed in this section is carried out by a SIB program written in Python. The Wi-Fi APs are uniformly deployed in a sensor field of 500×500 square meters with each AP having a transmission radius of 50 meters. The noise of the received RSSI signal is approximated as a white Gaussian noise $W\tilde{\mathrm{N}}(0, 0.7 \times d)$ from the experiment and d is the distance between the target device and AP. The target travels through the area by random walk with forwarding probability 0.6 and probability of going either left or right 0.2. The target walks 20 meters along the direction every time after the direction is selected.

In the SIBL algorithm, we initialize 2000 particles and the maximum number of iterations ranges from 1 to 5 in five different cases. The confidence bands of the target's trajectory are gradually improved from the increasing number of iterations being carried out in SIBL, and they are illustrated in Figures 2 to 5. The blue and red lines represent the trajectory of the moving target and the estimated path of the target respectively, while the purple ellipses are the confidence region for every estimated positions. The green triangles and circles represent the APs and their signal ranges.

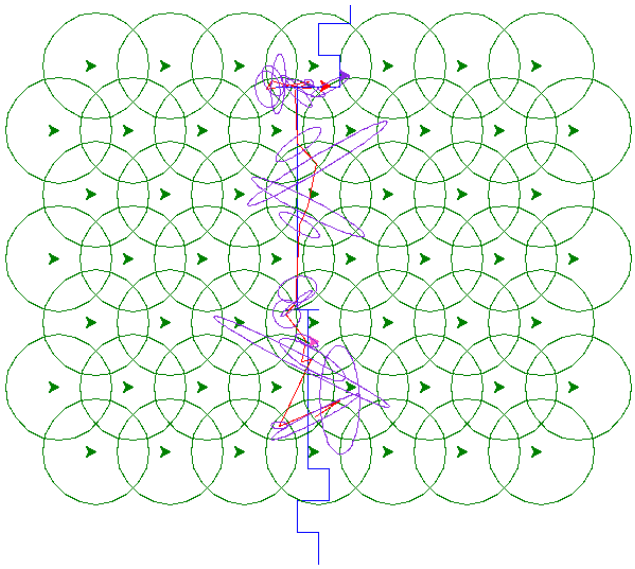

FIGURE 2. Confidence band with 1 iteration of SIBL.

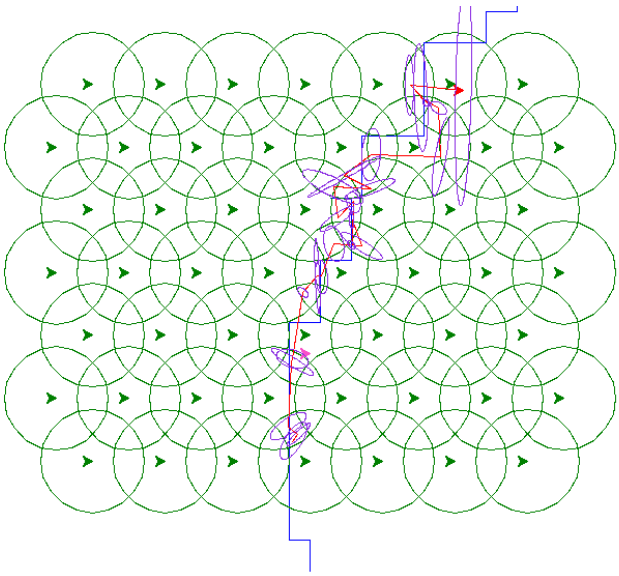

FIGURE 3. Confidence band with 2 iterations of SIBL.

In Figures 2 to 5, the confidence bands are shown at 95% confidence level for the moving trajectory. The completeness of these bands depend on the number of iterations executed in SIBL. It is obvious that the band constructed via few iterations is sparse and vice versa. With large number of iteration, the band begins to converge alongside with the path of the moving target. The positive dependency between the number of iteration and the computational time leads

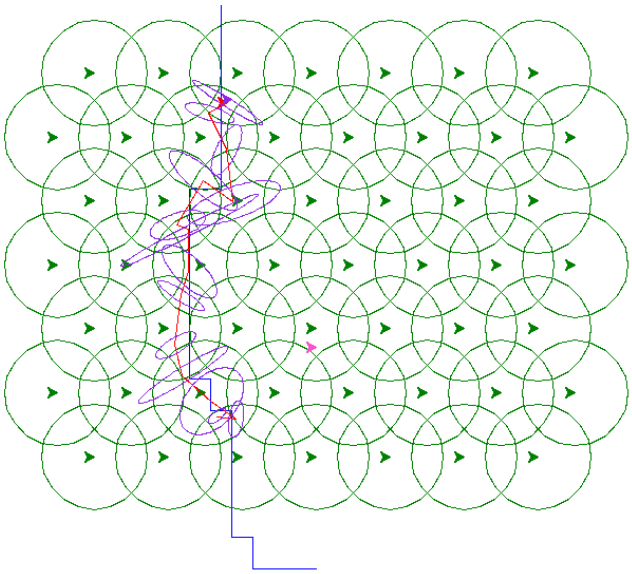

FIGURE 4. Confidence band with 3 iterations of SIBL.

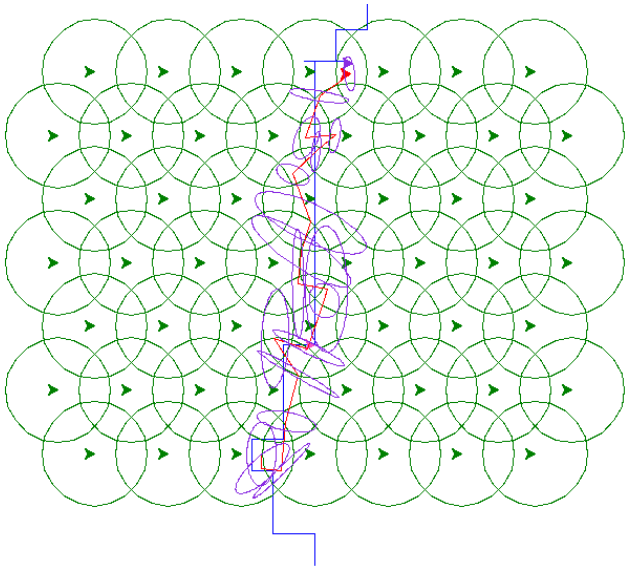

FIGURE 5. Confidence band with 5 iterations of SIBL.

to a tradeoff between the completeness and computational time of generating the confidence band. Nevertheless, most bands reach a well-converged state within 5 iterations in our simulation.

On the other hand, since the number of position estimations are related with the speed of the moving target, the effectiveness of confidence band for speed detection is demonstrated by comparing the average size of the confidence bands with different corresponding number of estimations. Larger amount of estimations represents a higher amount of speed. To ensure that the size of confidence band alters in the same fashion under different rate of RSSI estimation, the size of the confidence band with different corresponding number of estimations is also compared under various transmission frequency of PR. The simulation demonstrates the confidence band with the number of estimations within the range of 3 to 20 as benchmarks. The results are shown from Figures 6 to 9.

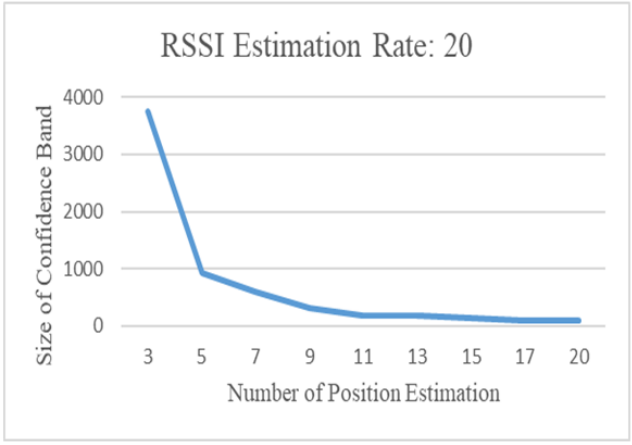

FIGURE 6. Band size under estimation rate 20.

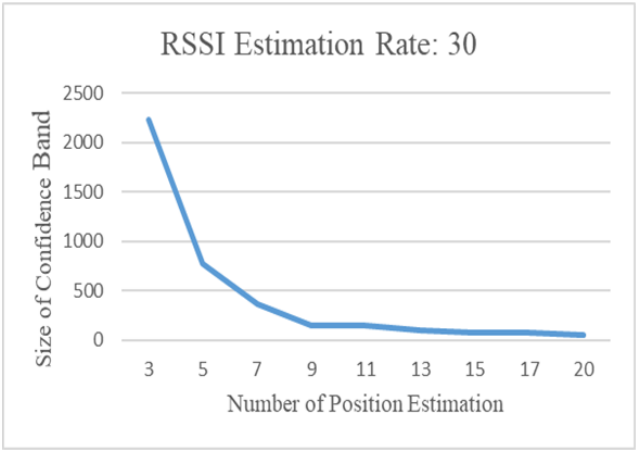

FIGURE 7. Band size under estimation rate 30.

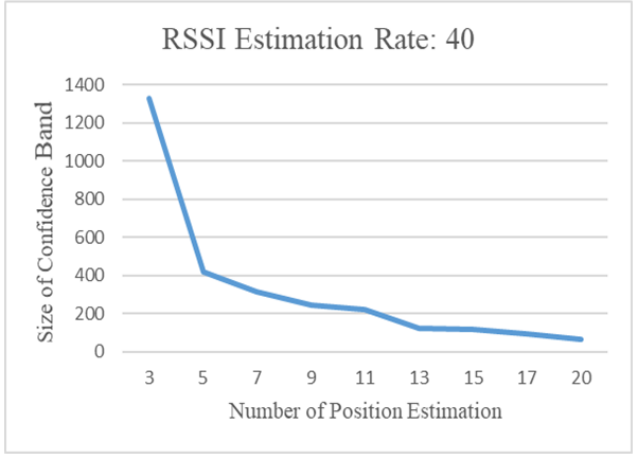

FIGURE 8. Band size under estimation rate 40.

The size of the confidence region is highly correlated to the number of position estimates under different rate of

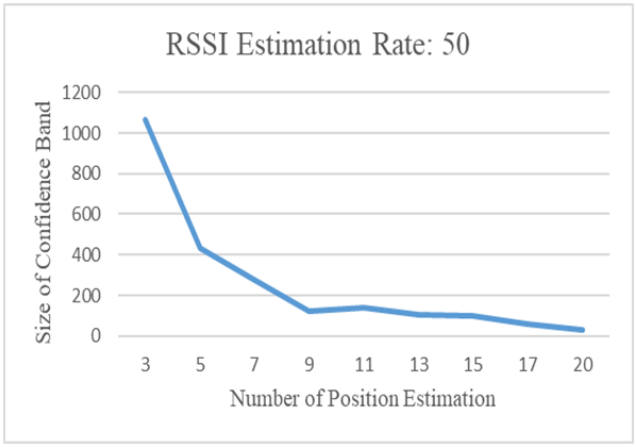

FIGURE 9. Band size under estimation rate 50.

RSSI estimation between the target and each APs. Under all different rates of RSSI estimation, the size of the confidence band decreases significantly as the number of localization increases from 3 to 9 and becomes relatively stable when the number of localization continues to increase. This size difference of the bands are remarkable among different speeds of the moving target, so it can clearly be used to reveal the speed of the target when a particular rate of RSSI estimation is set.

The rate of RSSI estimation is negatively related to the size of the confidence band as shown in Figure 10, but the alteration of the size between different frequencies at a fixed number of localization is not as notable as that between different amount of position estimations at a fixed rate of RSSI estimation. The speed of the moving target can be detected from the size of the band by knowing the number of RSSI estimation between the target and each AP. Thus, we can record in a database the size patterns of the confidence band with different number of estimation and different RSSI estimation frequency for data matching. The average speed of a target along the traveling trajectory can then be detected by matching the size of its corresponding confidence band in the database.

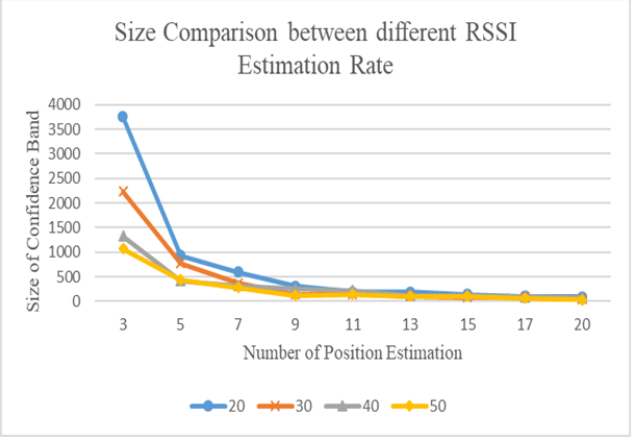

FIGURE 10. Band size under different estimation rate.

V. CONCLUSION AND FUTURE WORK

This paper proposes a new statistical method via a RSSI-based localization approach to construct real-time confidence sets and the confidence bands for the target's trajectory. We study the simultaneous inference on the estimated positions of the target and the dependency of the coordinates in each estimated positions. We also investigate the speed characteristics and patterns of the state of motion for a moving target that can be detected from the status of each confidence region in the constructed confidence band. To measure the positions adapted from the real environment, we propose a new objective function for localization. Then a modified SIB method called SIBL algorithm is then proposed for obtaining a confidence band with accuracy and speed. Experimental results shows that the confidence bands converge within a very small amount of iterations. Furthermore, the speed detection of a moving target is efficiently achieved without requiring any additional hardware implementations and large computational resources.

When we construct the confidence band, we apply the Bonferroni method to adjust M confidence regions. One may aware that this method is well-known to be conservative as the probability of simultaneous confidence band covers the corresponding true values for all collection of confidence sets is larger than $\frac{1-\alpha}{M}$. In many cases it goes to 1 when M is too large, resulting in a simultaneous confidence band that tends to be too wide and less informative. Thus, when M is very large, we suggest to consider some less conservative method on the adjustment of confidence regions. For example, it is possible to apply a precise simultaneous confidence bands for an unknown function introduced in [31].

To further reduce the computational costs, parallel computing architecture can be implemented to the SIBL algorithm with multi-core CPU, and it can be a promising extension from this work. In addition, we suspect that the transformation of the elements in each initialized particle of the SIBL into binary numbers may lead to a better chance to achieve global optimum before trapping in a local optimum. Furthermore, the concept of confidence band can be extended into prediction band, which can be used for characteristics predictive purposes. The target's path can then be predicted through such construction, and the target's position can also be revealed via the sampled datasets with prescribed probability. The target's future movement can be forecasted and it is of great interest in many traffic areas.

REFERENCES

[1] S. H. Wang, F. P.-C. Lin, and C.-P. Li, "Secure channel estimation method in TDD OFDM systems," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (IEEE BMSB)*, Jun. 2016, pp. 1–4, doi: 10.1109/BMSB.2016.7521979.2016.

[2] I. Vin, D. P. Gaillot, P. Laly, M. Liénard, and P. Degauque, "Overview of mobile localization techniques and performances of a novel fingerprinting-based method," *Comp. Rendus Phys.*, vol. 16, no. 9, pp. 862–873, 2015.

[3] W. W. Waqar, Y. Chen, and A. Vardy, "Smartphone positioning in sparse Wi-Fi environments," *Comput. Commun.*, vol. 73, pp. 108–117, Jan. 2016.

[4] X. Zhu and Y. Feng, "RSSI-based algorithm for indoor localization," *Commun. Netw.*, vol. 5, pp. 37–42, May 2013.

[5] S. Ravindra and S. N. Jagadeesha, "Time Of arrival based localization in wireless sensor networks: A linear approach," *Signal Image Process. Int. J.*, vol. 6, no. 1, pp. 45–49, 2015.
[6] S. Go and J.-W. Chong, "Improved TOA-based localization method with BS selection scheme for wireless sensor networks," *ETRI J.*, vol. 37, no. 4, pp. 707–716, 2015.
[7] N. H. Nguyen and K. Doğançay, "Optimal geometry analysis for multistatic TOA localization," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4180–4193, Aug. 2016.
[8] T.-K. Le and N. Ono, "Closed-form and near closed-form solutions for TOA-based joint source and sensor localization," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4751–4766, Sep. 2016.
[9] S. Zhang, S. Gao, G. Wang, and Y. Li, "Robust NLOS error mitigation method for TOA-based localization via second-order cone relaxation," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2210–2213, Dec. 2015.
[10] T. Qiao, Y. Zhang, and H. Liu, "Nonlinear expectation maximization estimator for TDOA localization," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 637–640, Dec. 2014.
[11] W. Jiang, C. Xu, L. Pei, and W. Yu, "Multidimensional scaling-based TDOA localization scheme using an auxiliary line," *IEEE Signal Process Lett.*, vol. 23, no. 4, pp. 546–550, Apr. 2016.
[12] Y.-H. Kim, D.-G. Kim, K.-H. Song, H.-N. Kim, and J. W. Han, "Analysis of sensor-emitter geometry for emitter localisation using TDOA and FDOA measurements," *IET Radar, Sonar Navigat.*, vol. 11, pp. 341–349, Feb. 2017.
[13] L. Yang and K. C. Ho, "An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4598–4615, Dec. 2009.
[14] T. Qiao, S. Redfield, A. Abbasi, Z. Su, and H. Liu, "Robust coarse position estimation for TDOA localization," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 623–626, Dec. 2013.
[15] S. Tomic, M. Beko, and R. Dinis, "Distributed RSS-AoA based localization with unknown transmit powers," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 392–395, Aug. 2016.
[16] S. Tomic, M. Beko, and R. Dinis, "3-D target localization in wireless sensor networks using RSS and AoA measurements," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3197–3210, Apr. 2017.
[17] A. Fascista, G. Ciccarese, G. Ricci, and A. Coluccia, "A localization algorithm based on V2I communications and AOA estimation," *IEEE Signal Process Lett.*, vol. 24, no. 1, pp. 126–130, Jan. 2017.
[18] P. Tarrío, A. M. Bernardos, and J. R. Casar, "Weighted least squares techniques for improved received signal strength based localization," *Sensors*, vol. 11, no. 9, pp. 8569–8592, 2011.
[19] F. P.-C. Lin and F. K. H. Phoa, "Performance study of parallel programming via CPU and GPU on swarm intelligence based evolutionary algorithm," in *Proc. ISMSI*, 2017, pp. 1–5.
[20] F. K. H. Phoa, "A Swarm Intelligence Based (SIB) method for optimization in designs of experiments," *Natural Comput.*, vol. 16, no. 4, pp. 597–605, 2017.
[21] F. K. H. Phoa and L. L. N. Chang, "A multi-objective implementation in swarm intelligence and its applications in designs of computer experiments," in *Proc. ICNC-FSKD*, Aug. 2016, pp. 253–258.
[22] F. K. H. Phoa, R.-B. Chen, W. Wang, and W. K. Wong, "Optimizing two-level supersaturated designs using swarm intelligence techniques," *Technometrics*, vol. 58, no. 1, pp. 43–49, 2016.
[23] M. Manaa and J. Akaichi, "Ontology-based modeling and querying of trajectory data," *Data Knowl. Eng.*, vol. 111, pp. 58–72, Sep. 2017.
[24] S. Kiatsupaibula, A. J. Hayterb, and S. Somsonga, "Confidence sets and confidence bands for a beta distribution with applications to credit risk management," *Insurance, Math. Econ.*, vol. 75, pp. 98–104, Jul. 2017.
[25] C. Chang, X. Lin, and R. T. Ogden, "Simultaneous confidence bands for functional regression models," *J. Stat. Planning Inf.*, vol. 188, pp. 67–81, Sep. 2017.
[26] S. Subramanian and P. Zhang, "Model-based confidence bands for survival functions," *J. Stat. Planning Inf.*, vol. 143, no. 7, pp. 1166–1185, 2013.
[27] P. Almers *et al.*, "Survey of channel and radio propagation models for wireless MIMO systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, p. 019070, Feb. 2007.
[28] B.-C. Liu, K.-H. Lin, and J.-C. Wu, "Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 499–509, Mar. 2006.
[29] W. Murphy and W. Hereman, "Determination of a position in three dimensions using trilateration and approximate distances," Dept. Math. Comput. Sci., Colorado School Mines, Golden, CO, USA, Tech. Rep. MCS-95-07, Oct. 1995.
[30] G. Cao, L. Wang, Y. Li, and L. Yang, "Oracle-efficient confidence envelopes for covariance functions in dense functional data," *Stat. Sinica*, vol. 26, pp. 359–383, Jan. 2016.
[31] J. Wang and L. Yang, "Polynomial spline confidence bands for regression curves," *Stat. Sinica*, vol. 19, pp. 325–342, Jan. 2009.

FRANK PO-CHEN LIN received the B.S. degree in electrical engineering from National Sun Yat-sen University, Taiwan, in 2016. He is currently pursuing the M.S. degree with the Institute of Communication Engineering, National Taiwan University, Taiwan. He is a Research Assistant with the Institute of Statistical Science, Academia Sinica. His current research interests include software defined networking, traffic engineering, parallel computing and wireless sensor networks. He was a recipient of the College Student Research Creativity Award in the Ministry of Science and Technology, Taiwan, in 2016.

FREDERICK KIN HING PHOA received the Ph.D. degrees in statistics from the University of California at Los Angeles (UCLA), Los Angeles, CA, USA, in 2009.

From 2009 to 2013, he was an Assistant Research Fellow with the Institute of Statistical Science, Academia Sinica, Taiwan, where he was promoted to an Associate Research Fellow, in 2013. He is an author of over 50 scientific articles, a speaker of over 90 invited talks in the international conferences and 60 seminar talks in the universities around the world. His research interests include design and analysis of physical, computer and network experiments, analysis of internet and social media data, network data analysis, nature-inspired metaheuristics optimization, big data analysis, stochastic control in large-scale systems, semiparametric methods to the data with missing covariates, deep learning and neural network modeling.

Dr. Phoa was a recipient of the Career Development Award in 2014, the Ta-You Wu Memorial Award (the Young Researcher Award) in 2014, and the Best Paper Award in the World Congress of Engineering in 2015. He was a recipient of the Special Talent Researcher Award from 2012 to 2017. He received the Excellent Young Researcher Project from 2013 to 2016 supported by the Ministry of Science and Technology (MOST), Taiwan, and the International Cost-Share Exchanges Scheme Project from 2016 to 2018 between the MOST and the Royal Society of U.K.