Time-of-Arrival Based Localization Under NLOS Conditions

Yiu-Tong Chan, *Senior Member, IEEE*, Wing-Yue Tsui, Hing-Cheung So, *Member, IEEE*, and Pak-chung Ching, *Senior Member, IEEE*

***Abstract*—Three or more base stations (BS) making time-of-arrival measurements of a signal from a mobile station (MS) can locate the MS. However, when some of the measurements are from non-line-of-sight (NLOS) paths, the location errors can be very large. This paper proposes a residual test (RT) that can simultaneously determine the number of line-of-sight (LOS) BS and identify them. Then, localization can proceed with only those LOS BS.**

The RT works on the principle that when all measurements are LOS, the normalized residuals have a central Chi-Square distribution, versus a noncentral distribution when there is NLOS. The residuals are the squared differences between the estimates and the true position. Normalization by their variances gives a unity variance to the resultant random variables.

In simulation studies, for the chosen geometry and NLOS and measurement noise errors, the RT can determine the correct number of LOS-BS over 90% of the time. For four or more BS, where there are at least three LOS-BS, the estimator has variances that are near the Cramer–Rao lower bound.

***Index Terms*—Mobile positioning, non-line-of-sight (NLOS), time-of-arrival (ToA).**

I. Introduction

DETERMATION of the position of a transmitting mobile station (MS) is a requirement in mobile phone operations. Responding to E-911 calls, providing location specific service information, and tracking a user, are examples that require an MS location [1].

An introduction to the basics of MS localization is given in [2]–[6]. Adding a GPS receiver to an MS, which appears to be an obvious solution, increases costs, weight, and power consumption, although in the U.S., this is now mandatory for new MS. In most other cases, the standard practice is to localize an MS from measurements of time-of-arrival (ToA), time-difference-of-arrival (TDoA), angle-of-arrival, signal strength, or a combination of these. The more popular is ToA. Three or more base stations (BS) measure the ToA of the transmission from an MS, giving ToA_i and $\delta_i = c \times \mathrm{ToA}_i$, which is the distance between the BS_i and the MS,. where the speed of light is c,. the locus of points at a distance δ_i from BS_i is a circle, and the intersections of circles give the MS location. Due to errors in ToA measurements, the circles do not intersect at a unique point, and it is necessary to find a location that best fits the measurements [1].

Another more significant problem is the non-line-of-sight (NLOS) situation, when the signal arrives at a BS from reflections. There is no direct, or line-of-sight (LOS), path. This often happens in an urban environment. Localization with an NLOS ToA can lead to large position errors.

There are, broadly, three ways to cope with the NLOS condition. The first way is similar to matched field processing for localization [7], [8]. It first measures the propagation characteristics of the channel, and then determines the MS location from a scattering model [9]–[12]. The difficulties are in obtaining an accurate model, and the model may change with the seasons and addition or removal of building structures.

The second way localizes with all NLOS and LOS measurements, but provides weighting or scaling to minimize the effects of the NLOS contributions. The weighting comes from either the localization geometry and BS layout [13], [14], or from the residuals (fitting errors) of individual BS [15]. The advantage here is that there is always an estimate, even when all BS are NLOS. The problem is that the answer can be unreliable because NLOS errors, though reduced, are always present.

The third way attempts to identify and localize with the LOS-BS. Identification is by a time-history based hypothesis test [16]; a probabilistic model [17]; [18]; residual information [19]; or maximum likelihood detection [25]. If the identification is correct, the accuracy is what the localization algorithm can provide, but there is always the possibility of wrong identification. In addition, unlike the first two methods, it requires at least three LOS-BS to localize.

This paper considers the third category. It determines the LOS dimension (D), i.e., the number of LOS-BS, identifies those BS, and localizes only with them. A residual test compares the residuals of a group of, say, k BS, against a predetermined threshold TH. If only a small percentage, say 10%, of the residuals are above TH, then $D = k$. Otherwise, the test moves to groups of $(k-1)$ BS. This process stops when it has determined a D, or when $D = 3$, the minimum value necessary for a unique localization. In the following, Section II describes the residual test, and Section III contains the simulation results. The conclusions are given in Section IV.

II. Residual Test

For clarity, "dimension" here in denotes the number of BS that receives LOS ToA measurements, and "identification" means finding the identities of those BS. The residual test performs

Manuscript received January 26, 2005; revised May 20, 2005. The review of this paper was coordinated by Dr. R. Klukas.

Y.-T. Chan is with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON, K7K 7B4, Canada (e-mail: chan-yt@rmc.ca).

P. Ching and W.-Y. Tsui are with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: pcching@ee.cuhk.edu.hk; wytsui@ee.cuhk.edu.hk).

H. C. So is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: hcso@ee.cityu.edu.hk).

Digital Object Identifier 10.1109/TVT.2005.861207

dimension determination and identification simultaneously but, conceivably, there can be schemes that separate the two functions.

A. ML Estimation

Let there be N LOS-BS at (x_i, y_i), measuring distance $\delta_i = c \times \mathrm{ToA}_i$ from an MS at unknown location $\Theta = [x\ y]^T$. Let

$$\delta_i = R_i + \varepsilon_i \tag{1}$$

where R_i are the true distances, and ε_i are independently and identically distributed (i.i.d.) zero mean Gaussian random variables denoting measurement noise. The maximum likelihood (ML) estimate is the Θ that maximizes the likelihood function

$$J = \sum_{i=1}^{N} (\delta_i - R_i)^2 \tag{2}$$

where

$$R_i^2 = (x - x_i)^2 + (y - y_i)^2. \tag{3}$$

Setting the differentials of (2) with respect to Θ to zero gives the likelihood equations

$$\sum_{i=1}^{N} \frac{(\delta_i - R_i)(x - x_i)}{R_i} = 0$$

$$\sum_{i=1}^{N} \frac{(\delta_i - R_i)(y - y_i)}{R_i} = 0. \tag{4}$$

The equations in (4) are nonlinear in Θ and have no closed form solutions. However, [20] was able to manipulate (4) into a set of linear equations in Θ, in the form of

$$A\Theta = b \tag{5}$$

but with A and b being functions of Θ. A suboptimal but linear algorithm [23] first gives an initial estimate of Θ, which can then give values of A and b. Solving (5) then produces a new value of Θ to update A and b, and then Θ. The procedure, called the approximate ML (AML) estimator [20], stops after five updates and takes the Θ that gives the smallest J in (2) as the solution. This ensures that the AML will not diverge, and will, at worst, have the errors of the linear estimator. Simulation results in [20] show that the AML actually attains the Cramer–Rao lower bound (CRLB).

When all measurements are LOS, the random variables in (7) should have a central χ^2 distribution. This implies that the estimates $\hat{x}_k$ and $\hat{y}_k$ should be Gaussian zero mean, and have variances equal to the CRLB.ML estimates obtained from the AML satisfy this requirement.

B. Dimension Determination

For ease of illustration but without loss of generality, let there be seven BS, with some or all δ_i being LOS. The problem is to determine the LOS dimension D and estimate Θ from the D number of BS.

The residual test (RT) begins by checking if $D = 7$. It computes, with the AML [20], a total of

$$\sum_{i=3}^{7} {}_7C_i = 99 \tag{6}$$

estimates of Θ. Thus, e.g., in the case of ${}_7C_3 = 35$, there are 35 different estimates of Θ, obtained from 7 δ_i, taken three at a time. Let these estimates be $\hat{\Theta}(k), k = 1 \ldots 99$, with $\hat{\Theta}(99)$ the ${}_7C_7$th estimate.

Next, the RT computes the square of the normalized residuals

$$\chi_x^2(k) = \frac{[\hat{x}(k) - \hat{x}(99)]^2}{B_x(k)}$$

$$\chi_y^2(k) = \frac{[\hat{y}(k) - \hat{y}(99)]^2}{B_y(k)}, \quad k = 1, \ldots 98. \tag{7}$$

The reference estimate is $\hat{\Theta}(99)$, which is the best estimate of the true Θ among all $\hat{\Theta}(k)$. This follows, because $\hat{\Theta}(99)$ is the estimate from all LOS (assumed) BS.

In (7), $B_x(k)$ and $B_y(k)$ are the approximation of the CRLB (see Appendix) of Θ from the δ_i combination that produces $\hat{\Theta}(k)$. The CRLB is the theoretical lower bound, valid when the ε_i in $\delta_i = R_i + \varepsilon_i$ are i.i.d. zero mean Gaussian random variables, and is a function of the localization geometry. Its computation requires knowledge of the true Θ. Since this is not available, the $\hat{\Theta}(k)$ location is used as a substitute to produce $B_x(k)$ and $B_y(k)$.

Now, if $D = 7$ and $\hat{\Theta}$ is an ML estimate of Θ, then

$$\frac{\hat{x}(k) - x}{\sqrt{B_x(k)}} \quad \text{and} \quad \frac{\hat{y}(k) - y}{\sqrt{B_y(k)}} \tag{8}$$

have a $N(0, 1)$ probability density function (pdf). It follows that the random variables (r.v.) in (7) have an approximate central χ^2 (Chi square) pdf of one degree of freedom [21]. If, however, one or more of the δ_i are NLOS, then $\hat{\Theta}(99)$ and some $\hat{\Theta}(k)$ will contain biases. The r.v. in (7) will have a noncentral χ^2 pdf, with the noncentral parameter being a function of the biases.

The pdf of the r.v. in (7) is only approximately central χ^2, even when $N = 7$, because 1) the $\hat{x}(k)$ and $\hat{y}(k)$ are partially correlated since they share some common ToA noise term ε_i in their computations; and 2) the CRLB in (7) are only approximations.

Fig. 1 plots the experimental pdfs of the r.v. $\chi_x^2(k)$ and $\chi_y^2(k)$ in (7), for $N = 7$, and $D = 7, D = 6$. It also contains the theoretical χ^2 pdf for comparison. The experimental pdfs come from histograms of 10 000 independent trials for noise $\sigma = 10$ m. For $N = 7$ and $D = 6$, the NLOS measurement has a bias $= 1000$ m. The x-axis in Fig. 1 begins at 0.1 so as to produce plots that can emphasize the differences between the theoretical and experimental pdfs. This figure shows that at $\chi^2 = 2.71$, the area to its right is 0.02 for the experimental pdf verses 0.07 for the theoretical [22].

Now, Fig. 1 suggests how the r.v. in (7) can determine the LOS dimension. For example, for a threshold $\mathrm{TH} = 2.71, D = 7$ if only 2% or less of the r.v. are larger than 2.71. Otherwise, the pdf is noncentral χ^2, and $D < 7$.

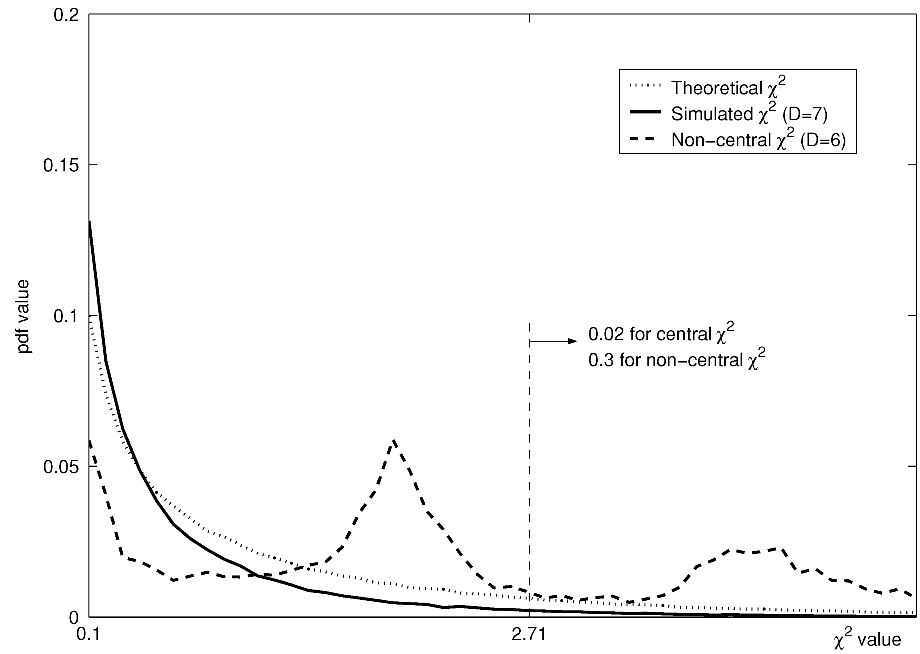


Fig. 1. Experimental and true χ^2 pdfs.

TABLE I
CONFUSION MATRIX FOR $\sigma = 9$ m

I \ D		Determined Dimension ($\hat{D}$) 7	6	5	4	3
True Dimension	7	943 / 943	53 / 53	4 / 4	0 / 0	0 / 0
	6	0 / 0	951 / 951	47 / 47	2 / 2	0 / 0
	5	0 / 0	0 / 0	948 / 948	51 / 51	1 / 1
	4	0 / 0	0 / 0	0 / 0	949 / 953	46 / 47
	3	0 / 0	0 / 0	0 / 0	0 / 8	853 / 992

TABLE II
CONFUSION MATRIX FOR $\sigma = 18$ m

I \ D		Determined Dimension ($\hat{D}$) 7	6	5	4	3
True Dimension	7	936 / 936	61 / 61	3 / 3	0 / 0	0 / 0
	6	0 / 0	943 / 943	55 / 55	2 / 2	0 / 0
	5	0 / 0	0 / 0	955 / 956	44 / 44	0 / 0
	4	0 / 0	0 / 0	0 / 0	934 / 942	47 / 58
	3	0 / 0	0 / 0	0 / 0	0 / 25	729 / 975

Let P_{OD} be the probability of overdetermination; i.e., when $D < 7$ but the RT decides that $D = 7$. Similarly, P_{UD} is the probability of underdetermination; i.e., when $D = 7$ but the decision is $D < 7$. For TH $= 2.71$, overdetermination occurs when $D < 7$ but only 2% or less of the r.v. are greater than 2.71. Underdetermination occurs when $D = 7$ but more than 2% of the r.v. are greater than 2.71.

It is desirable to minimize both P_{OD} and P_{UD} through a proper choice of TH. For TH < 0.1, $P_{OD} \simeq 1$ while $P_{UD} \simeq 0$, since both the χ^2 and noncentral χ^2 pdfs have approximately the same areas to the right of TH < 0.1. For TH > 0.1, $P_{OD} \simeq 0$ while P_{UD} is dependent on TH. For 10 000 independent trials, Fig. 2 gives a plot of P_{UD} verses TH. It shows that TH $= 2.71$ gives the lowest P_{UD}. The simulation experiments used TH $= 2.71$, resulting in almost zero overdetermination and a P_{UD} of less than 0.1, as seen in Tables I and II.

Now, in Fig. 1, the experimental χ^2 pdf has an area of 0.02 to the right of TH $= 2.71$. The theoretical value is actually 0.07. To accommodate other dimension tests besides seven, and to keep P_{OD} to a minimum instead of 2% or 7%, the RT checks to see if 10% of the r.v. in (7) are above TH for dimension determination. It is preferable to underdetermine than overdetermine. The former results in a smaller number of LOS BS for localization, while the latter contains NLOS BS for localization. resulting in large errors. Hence, for $N = 7$, if the number of r.v. larger than 2.71 is $0.1 \times 98 \times 2 \simeq 20$ or less, the RT decides that $D = 7$. Otherwise, it checks if $D = 6$.

Summarizing, the steps for deciding if $D = 7$ are as follows:

1) Do $\sum_{i=3}^{7} {}_7C_i$ estimates of $\hat{\Theta}(k), k = 1, \ldots, 99$.
2) Compute $B_x(k)$, and $B_y(k)$ and then $\chi_x^2(k)$ and $\chi_y^2(k)$ in (7).
3) Count the number of r.v., l, in (7) that are larger than TH $= 2.71$.
4) If $l \leq 20$, then $D = 7$, and $\hat{\Theta}(99)$ is the answer. Otherwise, check if $D = 6$.

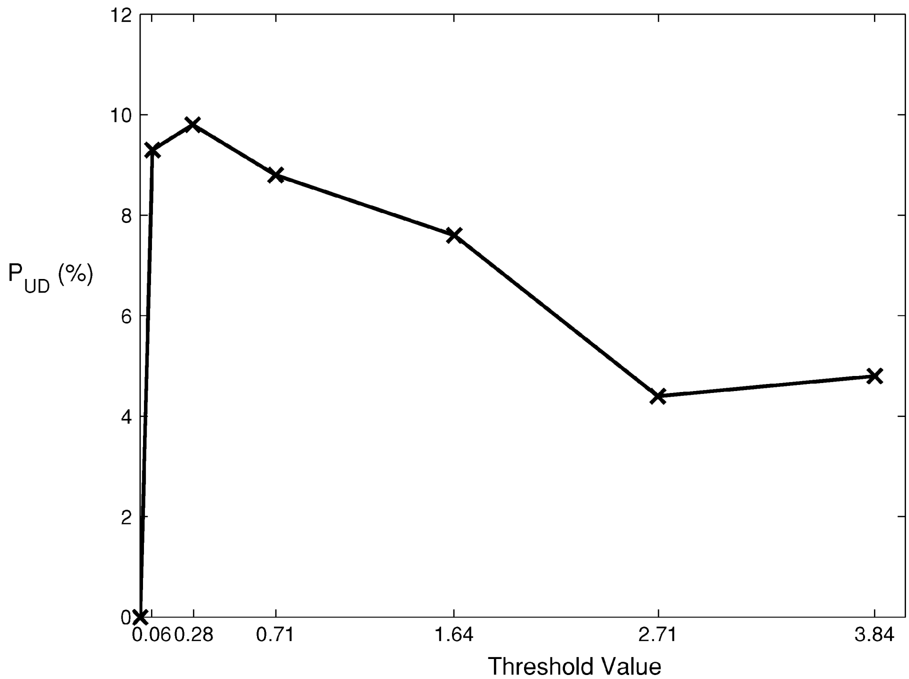


Fig. 2. Finding the TH.

To check $D = 6$, form ${}_7C_6 = 7$ sets of BS, six per set. For each set, do

$$\sum_{i=3}^{6} {}_6C_i = 42 \tag{9}$$

estimates of $\hat{\Theta}(k), k = 1, \ldots, 42$, and repeat steps 1) and 2), and then find seven l values. If there exists an $l \leq 0.1 \times 41 \times 2 \approx 8$, then $D = 6$, and that set of six BS are LOS BS. Otherwise, check for $D = 5$. If more than one set have $l \leq 8$, take the set with the smallest l, and the BS in this set are taken to be LOS. The steps for checking $D = 5$ and $D = 4$ are similar to the $D = 6$ case.

Now, suppose RT has determined that $D = 3$, so that the next task is to identify three LOS-BS from the seven. Since three BS do not provide a sufficient number of r.v. for a reliable RT, a different test is necessary. The following procedure, called the delta test (DT), takes two BS as the reference set, and combines with another BS out of the remaining five, to see if these three are LOS.

To illustrate, let the reference set contains BS_1 and BS_2, assumed to be LOS. Then, for $l = 1, 2$

$$\delta_l = R_l + \varepsilon_l \tag{10}$$

where the variables in (10) are defined as those in (1) and (3). In contrast, a NLOS BS_j will measure

$$\delta_j = R_j + \varepsilon_j + \alpha_j \tag{11}$$

where α_j is the additional distance due to NLOS. Letting

$$C^2 = x^2 + y^2 \tag{12}$$

and

$$K_i^2 = x_i^2 + y_i^2 \tag{13}$$

and expanding (10) and (11) result in

$$\begin{bmatrix} 2x_1 & 2y_1 & 0 \\ 2x_2 & 2y_2 & 0 \\ 2x_j & 2y_j & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \Delta_j \end{bmatrix} = \begin{bmatrix} K_1^2 + C^2 - \delta_1^2 \\ K_2^2 + C^2 - \delta_2^2 \\ K_j^2 + C^2 - \delta_j^2 \end{bmatrix} + \begin{bmatrix} 2R_1\varepsilon_1 + \varepsilon_1^2 \\ 2R_2\varepsilon_2 + \varepsilon_2^2 \\ 2R_j\varepsilon_j + \varepsilon_j^2 \end{bmatrix}. \tag{14}$$

In (14)

$$\Delta_j = 2R_j\alpha_j + 2\alpha_j\varepsilon_j + \alpha_j^2 \tag{15}$$

and the right-most vector contains unknown terms that are dependent on the r.v. ε_i.

Ignoring the vector of disturbance containing ε_i, solving (14) gives (x, y, Δ_j) in terms of C^2. Substituting the expressions for x and y into (12) results in a quadratic whose unknown is C^2. Its solution then gives numerical values for (x, y, Δ_j).

When all three BS in (14) are LOS, Δ_j should be $\simeq 0$, and nonzero when one or more of them are NLOS. The DT first forms ${}_7C_2 = 21$ reference sets of two BS per set, and computes a Δ_j from taking each of the other five BS as the jth BS. In those 21 sets, three BS will appear ${}_3C_2 = 3$ times, with one of the three taking turns as the jth BS. If these three BS are LOS, then $|\Delta_j|$ should be small. The DT sums the $3|\Delta_j|$ of these three BS and selects the set with the smallest sum as the LOS set.

Values of C^2 i.e., the roots of the quadratic, can be: 1) both equal and positive, meaning the solution for (x, y, Δ_j) is unique; 2) positive but unequal, and DT selects the smallest $|\Delta_j|$, since it is checking for LOS BS; and 3) both negative or complex, an indication that (14) does not have a real solution, and DT will consider this set of BS as containing at least one NLOS BS.

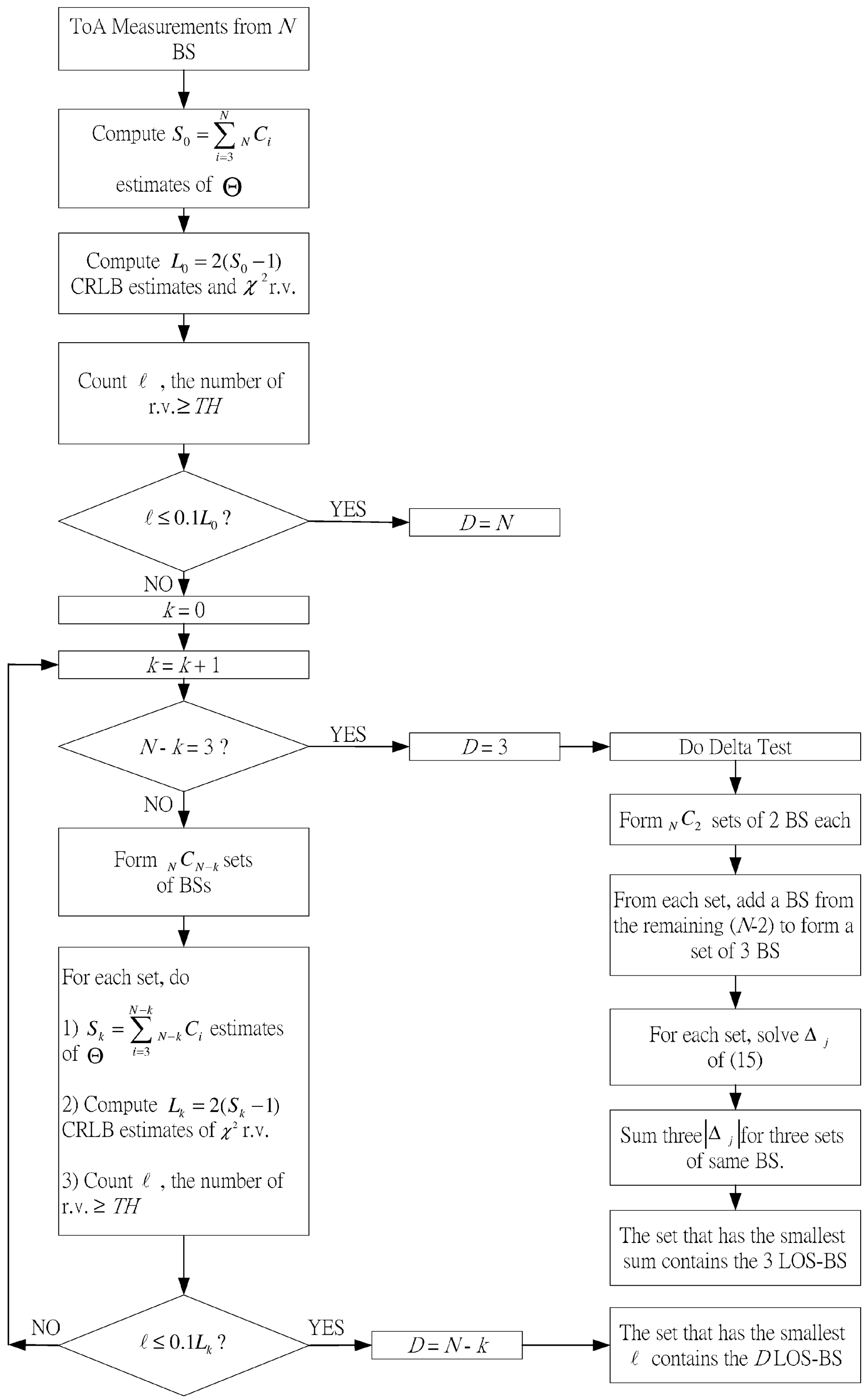


Fig. 3. Flowchart for the residual and delta tests.

To reiterate, the main idea of dimension determination is as follows. If there are N BS and the dimension is N, then the estimates of Θ from all possible BS combinations $\sum_{i=3}^{N} {}_N C_i$ should have normalized residuals that obey the central Chi square distribution. The algorithm flow chart is in Fig. 3.

III. Simulation Studies

The localization experiment in this section aims to evaluate the RT for dimension determination. There are seven BS, and their locations are at $(6000, 0)$, $(3000, -6000)$, $(-3000, -5000)$, $(-6000, -1000)$, $(-4000, 6000)$, $(0, 5000)$, and

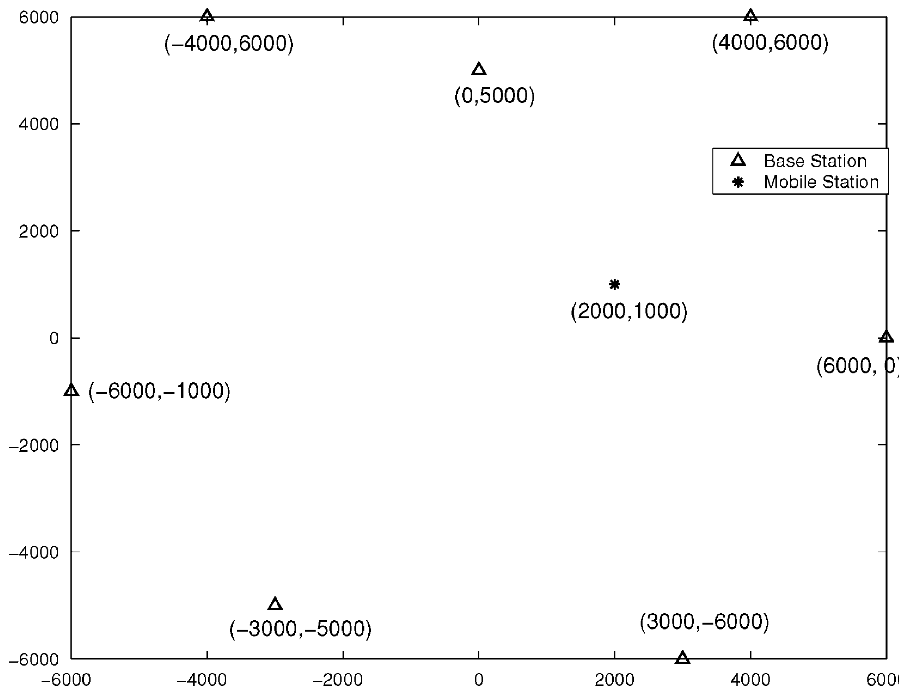


Fig. 4. Localization geometry.

$(4000, 6000)$, and the MS is at (2000, 1000). All units are in meters. Fig. 4 is the localization geometry. For a given trial, the distance measurements from ToA are, for LOS

$$\delta_i = R_i + \varepsilon_i \tag{16}$$

where ε_i are i.i.d. zero mean Gaussian r.v. of variance σ^2, and for NLOS

$$\delta_i = R_i + \varepsilon_i + \alpha_i \tag{17}$$

where α_i are i.i.d. r.v. of uniform probability density between 100 m and 1300 m. For a given dimension D and σ, there are 1000 independent trials. The NLOS and LOS BS are randomly selected for each trial. The RT determines $\hat{D}$ and gives the estimates.

The parameter D varies from three to seven, and σ from 0.05 m to 18.05 m. For wideband CDMA, the timing resolution is half a chip [26], equivalent to a distance of 40 m. Assuming a uniform pdf between 0 and 40 m, the ToA noise σ is $(40)/(\sqrt{3}) \simeq 23$ m. In GSM, a BS measures ToA in the uplink, with the 3G systems having an accuracy of about 10 m through GPS-assisted measurements [27]. Thus, the simulation range of σ is below that for WCDMA, but within those of 3G-GSM.

Tables I and II are the confusion matrices for $\sigma = 9$ m and $\sigma = 18$ m, respectively. There are two entries in each box in the tables. The lower entry gives the number of trials that have determined the dimension as $\hat{D}$. The upper gives the number of correct identifications (I) of the LOS BS. When $D \geq 5, I = \hat{D}$. But when D is four or three, it becomes more difficult to properly identify the LOS BS, and $I < \hat{D}$ in those cases.

As Tables I and II show, the RT provides at least 90% correct D determination. In addition, most of the incorrect $\hat{D}$ are smaller than D. When $\hat{D}$ is incorrect, it is preferable to have $\hat{D} < D$ rather than $\hat{D} > D$. The latter means the inclusion of NLOS measurements, which produces large estimation errors. In contrast, a $\hat{D} < D$ localization means not including some LOS BS. The degradation in accuracy is relatively minor compared to NLOS effects.

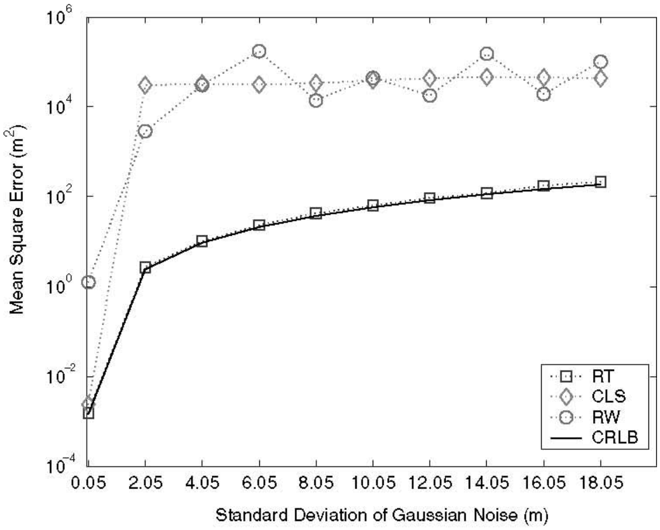
It is of value to compare the position estimation accuracy of RT with two other ToA-based NLOS mitigation methods, [15] and [24]. Fig. 5 plots the mean square errors (MSE) in m^2, from 1000 independent trials, as a function of the noise σ, where

$$\text{MSE} = \frac{\sum_{i=1}^{1000} \|\hat{\Theta}(i) - \Theta\|^2}{1000}. \tag{18}$$

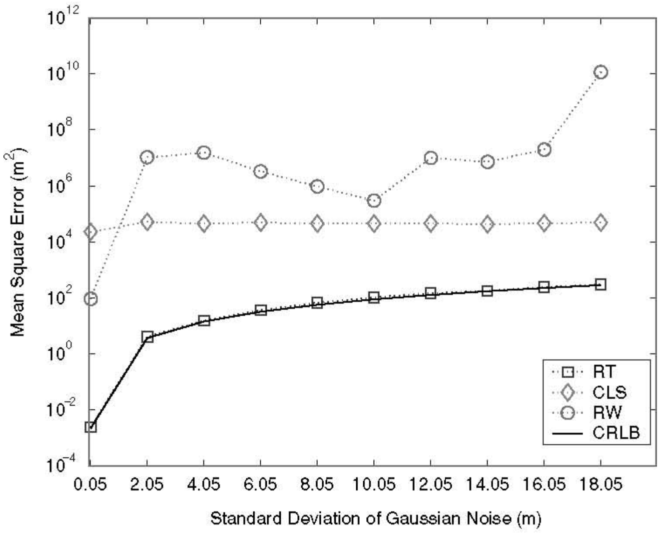
In (18), $\hat{\Theta}(i)$ is the estimate of the ith trial. The plots contain the MSE from RT, from the residual weighted (RW) estimate of [15], and from the constrained least squares (CLS) estimate of [24]. Also included as a reference is $10\log(\bar{B}_x + \bar{B}_y)$, where $\bar{B}_x$ and $\bar{B}_y$ are the averaged CRLB (See Appendix) for the estimation of x and y. Averaging over 1000 is necessary since for each trial, except for the $D = \text{seven}$ case, the set of LOS-BS can be different, resulting in a different CRLB.

In Fig. 5(a), where the true dimension $D = 7$, the MSE of RT follows closely, but slightly above, the CRLB. This is because, as seen in Tables I and II, some trials had determined a $\hat{D}$ lower than seven, and thus had less than seven BS for localization, resulting in a higher MSE. The MSE of RW is above the CRLB and the overall performance is inferior to RT, with much higher MSE at some σ. The CLS performed poorly in the situation when all BS are LOS. It assumes that all BS are NLOS, so that its MSE is quite large when this assumption is incorrect.

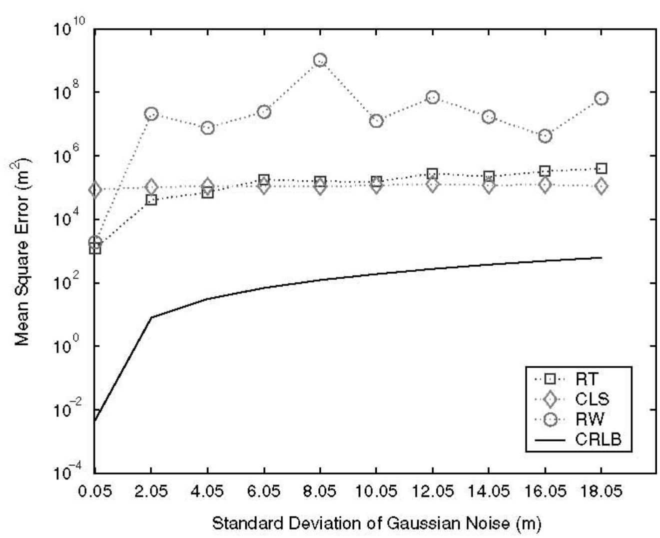
The results in Fig. 5(b), where $D = 5$, follow that of Fig. 5(a), except that RW has relatively even higher MSE. When $D = 3$, in Fig. 5(c), RT has a large percentage of incorrect LOS-BS



(a)



(b)



(c)

Fig. 5. (a) True dimension $= 7$. (b) True dimension $= 5$. (c) True dimension $= 3$.

identification. Its MSE, as well as that of RW and CLS, are all above the CRLB. Results for $D = 6$ and 4 are similar to the $D = 7$ and 5 cases, respectively, and are not shown.

To be fair to RW [15] and CLS [24], it should be explained that they both assume that all BS are NLOS, and do not select LOS-BS for localization. Hence, in a setting where there are LOS-BS, they do not perform as well as RT, which attempts to identify the LOS-BS.

IV. Conclusion

The localization of an MS can have significant errors when NLOS measurements are present. This paper has proposed a method, based on residual testing, to determine the LOS dimension and simultaneously identify the LOS BS. The principle is that if all measurements are LOS, and if the localization technique gives maximum likelihood estimates, then the residuals, normalized by the CRLB, will have a central χ^2 distribution. But if there are NLOS ToA, the distribution is noncentral χ^2. The check for noncentral χ^2 is by counting the number of normalized residuals that exceed a predefined TH value. If that number is larger than what a central χ^2 distribution should have, there is NLOS.

The RT method of LOS dimension determination is rather reliable, achieving accuracy of over 90% even when $\sigma = 18.05$ m. The location MSE is close to the CRLB for $D \geq 4$. But at $D = 3$, difficulty in identifying the LOS-BS leads to a higher MSE.

Appendix

CRLB

Let N be the number of LOS BS at locations (x_i, y_i) and let the MS be at (x, y). The CRLB is the theoretical lower bound on the estimation errors. It is a function of the localization geometry and the measurement noise σ^2. To find the CRLB, compute first the Fisher information matrix [23]. It is

$$I(\theta) = \frac{1}{\sigma^2}\begin{bmatrix} \sum \frac{(x-x_i)^2}{d_i^2} & \sum \frac{(x-x_i)(y-y_i)}{d_i^2} \\ \sum \frac{(x-x_i)(y-y_i)}{d_i^2} & \sum \frac{(y-y_i)^2}{d_i^2} \end{bmatrix}$$

where ${d_i}^2 = (x - x_i)^2 + (y - y_i)^2$ and the summation is from $i = 1$ to N. Then, the (1, 1) and (2, 2) elements of $I^{-1}(\theta)$ are, respectively, B_x and B_y, the CRLB for $\hat{x}$ and $\hat{y}$.

References

[1] J. J. Caffery, Jr., *Wireless Location in CDMA Cellular Radio Systems*. Norwell, MA: Kluwer, 1999.

[2] J. J. Caffery, Jr. and G. L. Stuber, "Overview of radiolocation in CDMA cellular systems," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 38–45, Apr. 1998.

[3] C. Drane, M. Macnaughton, and C. Scott, "Positioning GSM telephones," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 46–59, Apr. 1998.

[4] J. M. Zagami, S. A. Parl, J. J. Bussgang, and K. D. Melillo, "Providing universal location services using a wireless E911 location network," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 66–71, Apr. 1998.

[5] S. Tekinay, E. Chao, and R. Richton, "Performance benchmarking for wireless location systems," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 72–76, Apr. 1998.

[6] J. H. Reed, K. J. Krizman, B. D. Woerner, and T. S. Rappaport, "An overview of the challenges and progress in meeting the E-911 requirement for location service," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 30–37, Apr. 1998.

[7] D. F. Gingras, P. Gerstoft, N. L. Gerr, and C. F. Mecklenbrauker, "Electromagnetic matched field processing for source localization," in *Proc. Int. Conf. Acoustics Speech and Signal Proc., ICASSP-97*, vol. 1, Munich, Germany, Apr. 1997, pp. 479–482.
[8] Y. T. Chan, *Underwater Acoustics Data Processing*. Norwell, MA: Kluwer, 1988.
[9] R. B. Ertel and J. H. Reed, "Angle and time of arrival statistics for circular and elliptical scattering models," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1829–1840, Nov. 1999.
[10] S. Al-Jazzar, J. Caffery, Jr., and H. R. You, "A scattering model based approach to NLOS mitigation in TOA location systems," *Proc. IEEE Vehicular Technology Conf.*, vol. 2, pp. 861–865, May 2002.
[11] Radio Camera. The US Wireless Radio Camera network web page. 2000.
[12] S. Al-Jazzar and J. Caffery, Jr., "ML and Bayesian TOA location estimators for NLOS environments," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, Sep. 2002, pp. 1178–1181.
[13] S. Venkatraman, J. Caffery, Jr., and H. R. You, "Location Using LOS range estimation in NLOS environments," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, May 2002, pp. 856–860.
[14] N. Khajehnouri and A. H. Sayed, "A non-line-of-sight equalization scheme for wireless cellular location," in *Proc. ICASSP-03*, vol. 6, Apr. 2003, pp. 549–552.
[15] P. C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," in *Proc. IEEE Wireless Communications Networking Conf.*, 1999.
[16] M. P. Wylie and J. Holtzman, "The non-line-of-sight problem in mobile location estimation," in *Proc. IEEE Int. Conf. Universal Personal Communications*, vol. 2, 1996.
[17] J. Borras, P. Hatrack, and N. B. Mandayam, "Decision theoretic framework for NLOS identification," in *Proc. IEEE Vehicular Technology Conf.*, 1998.
[18] L. Xiong, "A selective model to suppress NLOS signals in angle-of-arrival (AOA)," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, 1998.
[19] L. Cong and W. Zhuang, "Non-line-of-sight error mitigation in TDOA mobile location," in *Proc. Global Telecommun. Conf.*, vol. 1, San Antonio, TX, Nov. 2001, pp. 680–684.
[20] Y. T. Chan, C. H. Yau, and P. C. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Trans. Veh. Technol.* to be published.
[21] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.
[22] E. S. Pearson and H. D. Hartley, *Biometrika Tables for Statisticians*. Cambridge, U.K.: Cambridge Univ. Press, 1966.
[23] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal. Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.
[24] X. Wang, Z. Wang, and B. O'Dea, "A ToA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 112–116, Jan. 2003.
[25] J. Riba and A. Urruela, "A non-line-of-sight mitigation technique based on ML-detection," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing ICASSP-04*, vol. 2, Montreal, QC, Canada, May 2004, pp. 153–156.
[26] Y. Zhao, "Standardization of mobile phone positioning for 3G systems," *IEEE Commun. Mag.*, vol. 40, no. 7, pp. 108–116, Jul. 2002.
[27] F. Gustafsson and F. Gunnarsson, "Possibilities and fundamental limitations of positioning using wireless communication networks measurements," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 41–53, Jul. 2005.

Yiu-Tong Chan (SM'80) was born in Hong Kong. He received the B.S. and M.S. degree from Queen's University, Kingtom, ON, Canada, and the Ph.D. degree from the University of New Brunswick, Moncton, NB, Canada.

He was an Engineer with Nortel Networks and has been a Professor in the Electrical and Computer Engineering Department at the Royal Military College of Canada, Kingston, serving as Head of the department from 1994 to 2000. From 2002 to 2005, he was a Visiting Professor at the Electronic Engineering Department of The Chinese University of Hong Kong. Presently, he is an Adjunct Professor at the Royal Military College. His research interests are in detection, estimation, localization, and tracking. He published a text entitled *Wavelet Basics* (Boston, MA: Kluwer, 1994).

Dr. Chan was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the Technical Chair of ICASSP-84, General Chair of ICASSP-91, Vice-Chair of ICASSP-03, and Social Chair of ICASSP-04. He directed a NATO ASI in 1988.

Wing-Yue Tsui was born in Hong Kong in 1978. She received the B.Eng. degree in electronic engineering from The Chinese University of Hong Kong, Shatin, NT, Hong Kong, in 2001.

From 2002 to 2003, she was a Research Assistant in the Computer Engineering and Information Technology Department of City University of Hong Kong. Since 2003, she has been working as a Research Assistant in the Electronic Engineering Department of The Chinese University of Hong Kong.

Hing-Cheung So (M'95) was born in Hong Kong. He received the B.Eng. degree from City University of Hong Kong and the Ph.D degree from The Chinese University of Hong Kong, Shatin, NT, in 1990 and 1995, respectively, both in electronic engineering.

From 1990 to 1991, he was an Electronic Engineer at the Research & Development Division of Everex Systems Engineering Ltd., Hong Kong. During 1995–1996, he worked as a Post-Doctoral Fellow at The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor at the Department of Electronic Engineering, City University of Hong Kong. Currently, he is an Associate Professor in the Department of Electronic Engineering at City University of Hong Kong. His research interests include adaptive filter theory, detection and estimation, wavelet transform, and signal processing for communications and multimedia.

Pak-chung Ching (SM'90) received the B. Eng. (First Class Hons.) and Ph.D. degrees from the University of Liverpool, U.K., in 1977 and 1981, respectively.

From 1981 to 1982 he was a Research Officer at the University of Bath, U.K. In 1982, he joined the Department of Electronic Engineering of the Hong Kong Polytechnic University as a Lecturer. Since 1984, he has been with the Department of Electronic Engineering of the Chinese University of Hong Kong (CUHK), Shatin, N.T., Hong Kong, where he is currently a Chair Professor. He was Department Chairman from 1995 to 1997, and was elected twice as Dean of Faculty of Engineering between 1997 and 2003. He has been appointed by CUHK as Head of Shaw College and Director of the Shun Hing Institute of Advanced Engineering in 2004. His research interests include digital signal processing, speech analysis/synthesis and recognition, adaptive filtering, array processing and digital communications.

Prof. Ching is active in participating professional activities. He is an Elected Council Member of the IEE. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1997 to 2000, and IEEE SIGNAL PROCESSING LETTERS from 2001 to 2003. He was a Member of the Technical Committee of the IEEE Signal Processing Society from 1996 to 2003. He is now the Chairman of the HKIE Publication Committee, and the Editor-in-Chief of the HKIE Transactions. He has been instrumental in organizing many international conferences in Hong Kong, including the 1997 IEEE International Symposium on Circuits and Systems, where he was the Vice-Chairman, and the IEEE International Conference on Acoustics, Speech and Signal Processing, where he was the Technical Program Chair. He received the IEEE Third Millennium Medal in 2000.