Localized Recursive Estimation in Energy Constrained Wireless Sensor Networks

Bang Wang, Kee Chaing Chua and Vikram Srinivasan
Department of Electrical and Computer Engineering
National University of Singapore
Email: {elewb, eleckc, elevs}@nus.edu.sg

Abstract— **This paper proposes a localized recursive estimation scheme for parameter estimation in wireless sensor networks. Given any parameter of a target occurring at some location and time, a number of sensors recursively estimate the parameter by using their local measurements of the parameter that is attenuated with the distance between a sensor and the target location and corrupted by noise. Compared with centralized estimation schemes that transmit all encoded measurements to a sink (or a fusion center), the recursive scheme needs only to transmit the final estimate to a sink. When the sink is faraway from the sensors and multihop communications have to be used, using localized recursive estimation can help to reduce energy consumption and reduce network traffic load. A sensor sequence with the fastest convergence rate is identified, by which the variance of estimation error reduces faster than all other sequences. In the case of adjustable transmission power, a heuristic has been proposed to find a sensor sequence with the minimum total transmission power when performing the recursive estimation. Numerical examples have been used to compare the performance of the proposed scheme with that of a centralized estimation scheme and have also shown the effectiveness of the proposed heuristic.**

Index Terms– **parameter estimation, recursive estimation, energy efficiency, wireless sensor networks.**

I. Introduction

Recently, *wireless sensor networks* (WSNs) that consist of a great number of sensors each capable of sensing, processing and transmitting environmental information have attracted a lot of research attentions [1]. Sensor networks can be used for example in environmental monitoring, military surveillance, space exploration etc. One of the important characteristics of sensor networks is that sensors cooperate to perform some functions. For example, sensors can cooperate to track an intruder or estimate a signal. This paper studies the estimation problem in a WSN: how to efficiently estimate a parameter of a target (e.g., the amplitude of a seismic or acoustic signal) located on a known location by cooperation of sensors.

Fig. 3 illustrate an example scenario for such target parameter estimation where several geographically distributed sensors are used to estimate the parameter of the target. The signal parameter might be attenuated with distance (as illustrated by the dashed concentric circles). Each sensor may make an local measurement of the attenuated parameter and the measure might be corrupted by noise. Since sensors are geographically distributed, their measurements of the same parameter thus have some spatial correlations. Furthermore, since each sensor is normally installed with a limited energy battery, energy consumption is a critical issue in sensor networks. Exploiting correlations in sensor networks can lead to significant potential advantages for the development of efficient communication protocols well-suited for the paradigm of WSNs [2]. The estimation efficiency here mainly refers to reducing energy consumption from the computations and transmissions when performing an estimation.

In WSNs, one way to reduce energy consumption is to use few bits to encode measurements while still meeting certain performance requirements [3]. Some schemes have been proposed for energy efficient estimation along this line [4][5][6][7] [8]. In these schemes, the basic idea is to reduce the bits needed to encode the local measurements. An universal distributed estimation scheme with one-bit message encoding functions has been proposed in [4][5] for a homogeneous WSN where sensors have observations of the same quality. In [6][7][8], an inhomogeneous sensing environment was considered, and the length of an encoded message is decided by the local signal to noise ratio. Suppose that we need K sensors' measurements to perform an estimation. These schemes require that all K sensors transmit their encoded measurements (messages) to a fusion center where an estimate is obtained via batch processing of the K messages. When the sink is far away from the K sensors and/or multi-hop communications have to be used, the transmission of these messages not only consumes much energy but also increases network traffic load even if the messages are short. Furthermore, it is generally believed that in a WSN the energy required for local computation is much less than that used for communications [1][9].

The above observations motivate us to propose a localized recursive estimation scheme in this paper. The proposed scheme performs the estimation locally and recursively by the sensors measuring the parameter. A sensor performs estimation based on its own measurement and the intermediate estimation results from its upstream sensor. After performing estimation, it then transmits its intermediate estimation results to the next downstream sensor. At the termination of estimation, the last sensor transmits only the final estimate to the sink. In a large scale WSN, the communications overhead hence can be

[0] This paper was presented in part at the International Conference on Mobile Ad-hoc and Sensor Networks (MSN05), Wuhan, China, 2005.

reduced significantly. Furthermore, unlike the measurement model used in [4][5][6][7] [8], we use a more realistic distance based measurement model where signals decay with distance when a wavefront expands [9]. Based on this measurement model, a sensor sequence with the fastest convergence rate is identified and by which the variance of estimation error reduces faster than all other sequences. In the case of adjustable transmission power, we propose a heuristic to schedule sensors' transmissions in order to further reduce the total transmission power for the sensors transmitting their intermediate estimation results. Numerical examples have been used to compare the performance of the proposed scheme with that of a centralized estimation scheme and have also shown the effectiveness of the proposed heuristic.

The rest of the paper is organized as follows. Section II presents the problem formulation for parameter estimation in sensor networks. A localized recursive parameter estimation scheme is provided and its property is analyzed in Section III. Section IV gives some numerical examples and some concluding remarks are provided in Section VI.

II. Problem Formulation

Consider a snapshot of a sensor field with a set of K geographically distributed sensors, each making a measurement on an unknown parameter θ at some location and time. The scenario can be the sensing an acoustic or seismic signal of amplitude θ. Let $d_k, k = 1, 2, ..., K$ denote the distance between a sensor k and the parameter θ. The parameter θ is assumed to decay with the distance, and at distance d it is θ/d^α, where $\alpha > 0$ is the decay component. Furthermore, the measurement of the parameter at a sensor is corrupted by an additive noise. This model has been widely used to model acoustic signals [10] and applied for acoustic source localization [11]. Let x_k denote the measurement of the parameter θ at a sensor k, and is given by

$$x_k = \frac{\theta}{d_k^\alpha} + n_k, k = 1, 2, ..., K. \quad (1)$$

The objective of a parameter estimator is to estimate θ based on the corrupted measurements. Let $\hat{\theta}$ and $\tilde{\theta} = \theta - \hat{\theta}$ denote the estimate and the estimate error, respectively. A commonly used criterion is to minimize the *mean squared error* (MSE) of an estimator, i.e., to minimize $\mathbb{E}[\tilde{\theta}^2]$. Since sensors are geographically distributed, they need to encode the measurements to messages and then these messages are transmitted to a fusion center which generate an estimate according to a fusion function. The design of these message functions ($\{m_k : k = 1, 2, ..., K\}$) and the fusion function ($\Gamma(m_1, m_2, ..., m_K)$) has been discussed in [4][8], where an estimate is made based on a batch process of all the measurements. We note that these schemes are referred to as distributed estimation schemes in [4][8] as the measurements are distributed. However, the estimation processing (the fusion function) is centralized and is based on the availability of the batch measurements, which needs all these K measurements to be transmitted to the center. Hence, their estimation scheme is called *centralized batch estimation scheme* (CBES) in this paper, and is illustrated by Fig. 1.

It is generally believed that in sensor networks the energy required for local computation is much less than that used for communications. In CBES, transmitting all measurements to the fusion center might cause high traffic load and consume much energy in a WSN, especially when the fusion center is far away where multi-hop communications are needed. This motivates us to propose to use a *localized recursive estimation scheme* (LRES) to address this problem. In LRES, K sensors are considered to be connected in tandem (via radio) for parameter estimation. Without loss of generality, assume sensor k is a predecessor of sensor $k+1$. This means sensor $k+1$ makes an estimate based on its own measurements x_k and its predecessor's output. Finally, the sensor K makes the last estimate and sends it to the sink. Suppose that each sensor is i hops away from the sink while any two successive sensors are only one hop away, a simple calculation shows a total of a number of $(i-1) \times (K-1)$ communications can be saved. Furthermore, if an estimate is suitably transmitted, a sensor no longer needs a message function to encode x_k to a message m_k, which further reduces energy consumption. An LRES is illustrated by Fig. 2. We note that to estimate $\hat{\theta}(k)$, another parameter $B(k-1)$ together with the estimate $\hat{\theta}(k-1)$ from its predecessor sensor need to be transmitted. The definition and computation of $B(k)$ is given in next section. Compared with the CBES illustrated by Fig. 1, the LRES need not encode the measurements, however, the parameter $B(k)$ and the estimate $\hat{\theta}(k)$ might need to be suitably encoded and transmitted. The next section presents a general recursive estimation algorithm to construct a local fusion function based on a best linear unbiased estimator.

Some notation remarks are summarized here. We use $\hat{\theta}_k$ to denote the estimate from a CBES when k measurements are used for estimation, and $\hat{\theta}(k)$ to denote the estimate produced by the kth sensor in LRES after the estimates from its previous $k-1$ sensors. Boldface is used to denote vectors and matrixes; and $\mathbf{D}^T$ denotes the transpose of matrix $\mathbf{D}$.

III. A Localized Recursive Estimation Scheme

A. A recursive estimation algorithm

The measurement given by (1) can be written in matrix form for K sensors as

$$\mathbf{X} = \mathbf{D}\theta + \mathbf{N}, \quad (2)$$

where $\mathbf{X} = (x_1, x_2, ..., x_K)^T$, $\mathbf{D} = (d_1^{-\alpha}, d_2^{-\alpha}, ..., d_K^{-\alpha})^T$, and $\mathbf{N} = (n_1, n_2, ..., n_K)^T$. The additive noises are assumed to be spatially uncorrelated white noise with zero mean and σ_k^2 variance, but otherwise unknown. The covariance matrix of noises $\{n_k : 1, 2, ..., K\}$ is given by

$$\mathbf{R} = \mathbb{E}[\mathbf{N}\mathbf{N}^T] = \mathrm{diag}[\sigma_1^2, \sigma_2^2, ..., \sigma_K^2]. \quad (3)$$

Note that $\mathbf{R}$ is symmetric and positive definite. A well-known *best linear unbiased estimator* (BLUE) [12] can be

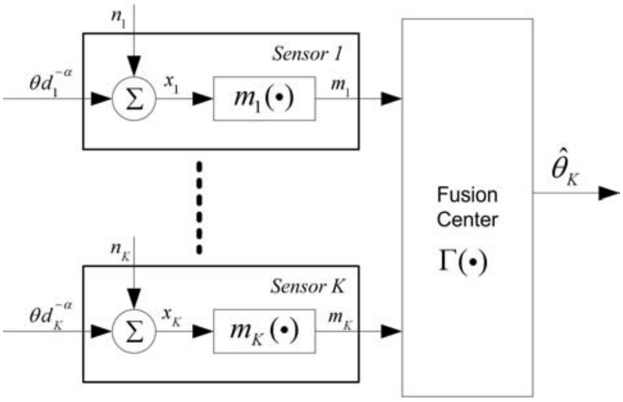

Fig. 1. Illustration of a centralized batch estimation scheme (CBES).

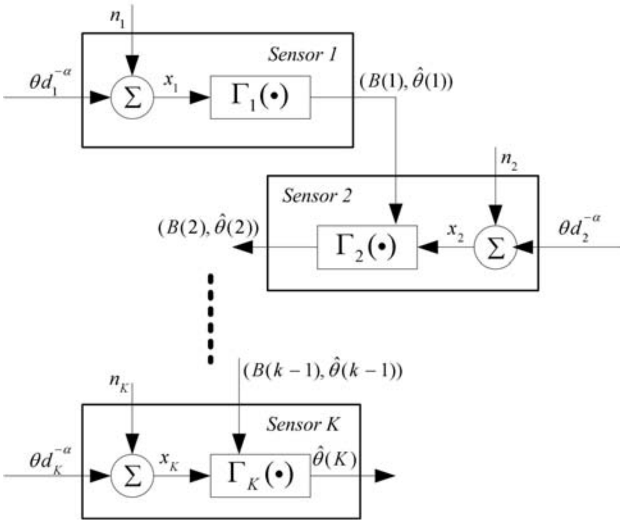

Fig. 2. Illustration of a localized recursive estimation scheme (LRES).

applied to estimate $\hat{\theta}_K$ and to achieve a minimum MSE. According to BLUE, when K measurements are available, the estimate $\hat{\theta}_K$ of the original signal θ is given as

$$\hat{\theta}_K = \left[\mathbf{D}^T\mathbf{R}^{-1}\mathbf{D}\right]^{-1}\mathbf{D}^T\mathbf{R}^{-1}\mathbf{X}. \quad (4)$$

The MSE of BLUE is given as

$$\mathbb{E}[(\theta - \hat{\theta}_K)^2] = (\mathbf{D}^T\mathbf{R}^{-1}\mathbf{D})^{-1}, \quad (5)$$

and the estimation error $\tilde{\theta}$ is given as

$$\tilde{\theta}_K = \theta - \hat{\theta}_K = -[\mathbf{D}^T\mathbf{R}^{-1}\mathbf{D}]^{-1}\mathbf{D}^T\mathbf{R}^{-1}\mathbf{N}. \quad (6)$$

Note that Eq.(4) also gives the fusion function Γ_K when K measurements are available. The following theorem provides a recursive BLUE for our problem.

Theorem 1: The kth sensor can make an estimate $\hat{\theta}(k)$ by using the following recursive structure

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{B(k)}{d_k^\alpha\sigma_k^2}\left(x_k - \frac{\hat{\theta}(k-1)}{d_k^\alpha}\right), k = 1, 2, ..., K, \quad (7)$$

where

$$B(k) = \left(\frac{1}{B(k-1)} + \frac{1}{d_k^\alpha\sigma_k^2}\right)^{-1} \quad (8)$$

These equations are initialized by $\hat{\theta}(0) = 0$ and $B(0)$ equal to a very large number.

Proof: With a little abuse of notation, we use subscript k to indicate the dimension for vectors and matrix. For example, $\mathbf{X}_k = (x_1, x_2, ..., x_k)^T$ denotes a $k \times 1$ vector, and $\mathbf{R}_k$ denotes a $k \times k$ matrix. The objective of our recursive estimator is to produce the same estimate as that of the batch estimator. This implies $\hat{\theta}_k = \hat{\theta}(k)$, $k = 1, 2, ..., K$. Hence we rewrite (4) as

$$\hat{\theta}(k) = \hat{\theta}_k = [\mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{D}_k]^{-1}\mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{X}_k. = B(k)A(k) \quad (9)$$

where $A(k) = \mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{X}_k$ and $B(k) = [\mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{D}_k]^{-1}$. With some algebra, we have

$$\begin{aligned} A(k) &= \mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{X}_k \\ &= \sum_{i=1}^{k}\frac{x_i}{d_i^\alpha\sigma_i^{-2}} = \sum_{i=1}^{k-1}\frac{x_i}{d_i^\alpha\sigma_i^2} + \frac{x_k}{d_k^\alpha\sigma_k^2} \\ &= A(k-1) + \frac{x_k}{d_k^\alpha\sigma_k^2}. \end{aligned} \quad (10)$$

and

$$\begin{aligned} B^{-1}(k) &= \mathbf{D}_k^T\mathbf{R}_k^{-1}\mathbf{D}_k \\ &= \sum_{i=1}^{k}\frac{1}{d_i^{2\alpha}\sigma_i^2} = \sum_{i=1}^{k-1}\frac{1}{d_i^{2\alpha}\sigma_i^2} + \frac{1}{d_k^{2\alpha}\sigma_k^2} \\ &= B^{-1}(k-1) + \frac{1}{d_k^{2\alpha}\sigma_k^2} \end{aligned} \quad (11)$$

Rewrite (11) as

$$B(k) = \left(\frac{1}{B(k-1)} + \frac{1}{d_k^{2\alpha}\sigma_k^2}\right)^{-1} \quad (12)$$

Now we come to derive the relationship between $\hat{\theta}(k)$ and $\hat{\theta}(k-1)$. By using (10), (11) and (12), and the fact $A(k-1) = B^{-1}(k-1)\hat{\theta}(k-1)$, we obtain

$$\begin{aligned} \hat{\theta}(k) &= B(k)A(k) = B(k)\left[A(k-1) + \frac{x_k}{d_k^\alpha\sigma_k^2}\right] \\ &= B(k)\left[(B^{-1}(k) - \frac{1}{d_k^{2\alpha}\sigma_k^2})\hat{\theta}(k-1) + \frac{x_k}{d_k^\alpha\sigma_k^2}\right] \\ &= \hat{\theta}(k-1) + \frac{B(k)}{d_k^\alpha\sigma_k^2}\left(x_k - \frac{\hat{\theta}(k-1)}{d_k^\alpha}\right) \end{aligned} \quad (13)$$

Note that when choosing $\hat{\theta}(0) = 0$ and $B(0)$ equal to a very large number, $B(1) \approx d_1^{2\alpha}\sigma_k^2$ and hence

$$\hat{\theta}(1) = d_1^\alpha x_1 = \theta + d_1^\alpha n_1 \quad (14)$$

which is the same as $\hat{\theta}_1$, i.e., set $K = 1$ in (4). ■

The first sensor only needs its local measurement x_1, its distance information d_1^α and its local noise variance σ_1^2 to compute $\hat{\theta}(1)$. For kth sensor ($k = 2, 3, ...K$), the computation of $\hat{\theta}(k)$ needs the estimate and $B(k-1)$ from its predecessor sensor $\hat{\theta}(k-1)$ as well as its local measurement, its distance information and its noise variance. Since $B(k-1)$ is a scalar, the transmission overhead for $B(k-1)$ is small. In (7), the term $\hat{\theta}(k-1)/d_k^\alpha$ can be considered as a prediction of the actual measurement of x_k, and hence $[x_k - \hat{\theta}(k-1)]/d_k^\alpha$ can be considered as the prediction error. Consequently, $\hat{\theta}(k)$ can be considered as a combination of the just computed $\hat{\theta}(k-1)$ with a linear transformation of the prediction error.

Since $\hat{\theta}(k) = \hat{\theta}_k$ for $k = 1, 2, ..., K$, some properties of the batch estimator (4) are also applicable to the recursive estimator (7). Recall that the estimator given by (4) is an unbiased estimator since $\mathbb{E}[\hat{\theta}_K] = \theta$. Accordingly, the following unbiasedness definition is used for the recursive estimator. Furthermore, the recursive estimator (7) is also an unbiased estimator.

Definition 1: A recursive estimator $\hat{\theta}(k)$ is an unbiased recursive estimator of a deterministic parameter θ if

$$\mathbb{E}[\hat{\theta}(k)] = \theta, \text{for all } \ k = 1, 2, ..., K. \quad (15)$$

Corollary 1: The recursive estimator given by (7) is an unbiased recursive estimator.
The proof can be based on the fact that $\hat{\theta}(k) = \hat{\theta}_k$ or by a simple induction argument. From Corollary 1, the variance of the kth estimate $\mathbb{V}[\hat{\theta}(k)] = \mathbb{E}[(\hat{\theta}(k)-\theta)^2]$ and can be obtained via (5). Furthermore, $\mathbb{V}[\hat{\theta}(k)] = B(k)$ and (8) can be considered as a recursive structure of $\mathbb{V}[\hat{\theta}(k)]$.

Corollary 2: The variance of the kth estimate $\mathbb{V}[\hat{\theta}(k)]$ of the recursive estimator satisfies

$$\begin{aligned} \mathbb{V}[\hat{\theta}(k)] &= \mathbb{E}[(\hat{\theta}(k) - \theta)^2] \\ &= \left(\frac{1}{\mathbb{V}[\hat{\theta}(k-1)]} + \frac{1}{d_k^{2\alpha}\sigma_k^2} \right)^{-1} \end{aligned} \quad (16)$$

and

$$\mathbb{V}[\hat{\theta}(k)] < \mathbb{V}[\hat{\theta}(k-1)] \quad (17)$$

for $k = 1, 2, ..., K$, where $1/\mathbb{V}[\hat{\theta}(0)] \approx 0$ by appropriately choosing $\mathbb{V}[\hat{\theta}(0)]$.
Since the recursive estimator computes the variance recursively, the recursive estimator is also called the *variance form* of recursive BLUE. Furthermore, (17) indicates that the more sensors are used for parameter estimation, the better the estimation performance in terms of mean squared estimation error.

The recursive estimator can stop the estimation when the error performance is achieved or when the improvement of the estimate is small. Let ϵ denote the required error variation. When $B(k) \le \epsilon$, the kth sensor stops forwarding the estimate $\hat{\theta}(k)$ and $B(k)$ to its successor but sends to the sink. Another criterion is that when $B(k) - B(k-1) \le \epsilon'$, the recursive estimation stops.

B. The sensor sequence with the fastest convergence rate

The choice of the order of the sensors to perform parameter estimation impacts the convergence rate (to a given error threshold) of the recursive estimator. In some cases, we might not use all the K sensors when we already obtain an estimate with required error performance. Corollary 1 states that the recursive estimator is an unbiased one and the mean of the estimation error $\mathbb{E}[(\theta - \hat{\theta}(k))] = 0$ for all k. However, the variance of the estimation error of the recursive estimator might be dependent on the choice of the sequence of sensors as well as the number of measurements. Recall that BLUE is designed for minimizing MSE. This motivates us to use the variation of the estimation error to measure the convergence rate. From Corollary 2, this is equal to using $\mathbb{V}[\hat{\theta}(k)]$ to measure the estimation efficiency. Let $\overline{K} =< k_1, k_2, ..., k_K >$ denote a sequence of K sensors. The following definition gives efficiency measurement for choices of the order of sensors, and Theorem 2 provides necessary conditions for a sequence with the fastest convergence rate.

Definition 2: A sequence of sensors $\overline{K} =< k_1, k_2, ..., k_K >$ is said to converge faster than another sequence $\overline{K}' =< k'_1, k'_2, ..., k'_K >$ if

$$\mathbb{V}[\hat{\theta}(k_i)] \le \mathbb{V}[\hat{\theta}(k'_i)], \text{ for all } i = 1, 2, ..., K. \quad (18)$$

Theorem 2: The sensor sequence $\overline{K} =< k_1, k_2, ..., k_K >$ with the fastest convergence rate should satisfy

$$d_{k_1}^{2\alpha}\sigma_{k_1}^2 \le d_{k_2}^{2\alpha}\sigma_{k_2}^2 \le \cdots \le d_{k_K}^{2\alpha}\sigma_{k_K}^2. \quad (19)$$

Proof: The proof proceeds by induction on the construction of $\overline{K}$. For $i = 1$, we have $\hat{\theta}(1) = \theta + d_1^\alpha n_1$ and hence

$$\mathbb{V}[\hat{\theta}(k_1)] = \mathbb{E}[(\hat{\theta}(k_1) - \theta)^2] = d_{k_1}^{2\alpha}\sigma_{k_1}^2. \quad (20)$$

Obviously the smallest $\mathbb{V}[\hat{\theta}(k_1)]$ is achieved when choosing the sensor k_1 such that $d_{k_1}^{2\alpha}\sigma_{k_1}^2$ is the smallest among all sensors. Now assume the construction of $\overline{K}$ is the most efficient for i sensors, i.e., the sequence of selected sensors is as $< k_1, k_2, ..., k_i >$ such that $d_{k_1}^{2\alpha}\sigma_{k_1}^2 \le d_{k_2}^{2\alpha}\sigma_{k_2}^2 \le ... \le d_{k_i}^{2\alpha}\sigma_{k_i}^2$ and $\mathbb{V}[\hat{\theta}(k_i)]$ is the one with the fastest convergence rate for $1, ..., i$. From (16), we have

$$\mathbb{V}[\hat{\theta}(k_{i+1})] = \frac{1}{\frac{1}{\mathbb{V}[\hat{\theta}(k_i)]} + \frac{1}{d_{k_{i+1}}^{2\alpha}\sigma_{k_{i+1}}^2}} \quad (21)$$

Consider another sequence construction $\overline{K}'$ which is also with the fastest convergence rate for the first i sensors, however, the $(i+1)$th sensor is different from $\overline{K}$. From (21), it is easily to see that $\mathbb{V}[\hat{\theta}(k_{i+1})] \le V[\hat{\theta}(k_{i+1})]$ implies that $d_{k_{i+1}}^{2\alpha}\sigma_{k_{i+1}}^2 \le d_{k'_{i+1}}^{2\alpha}\sigma_{k'_{i+1}}^2$, that is, the selection of the k_{i+1} sensor should also satisfy (19). Hence the desired result is obtained from the induction. ■
When all sensors have the same noise variance, we have the following corollary.

Corollary 3: When all sensors have the same noise variance, i.e., $\sigma_1^2 = \sigma_2^2 = ... = \sigma_K^2$, the most efficient

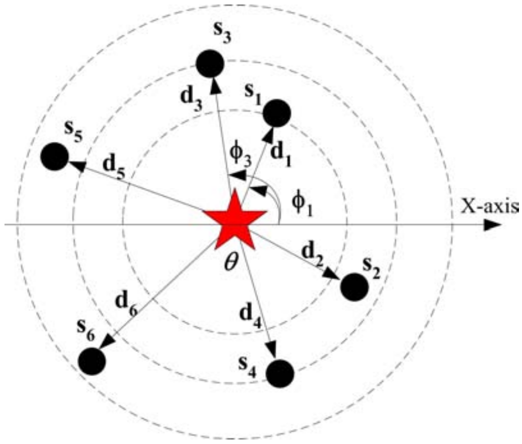

Fig. 3. Target Parameter Estimation by Geographically Distributed Cooperative Sensors. The parameter is attenuated with distances.

sequence of sensors $\overline{K} =< k_1, k_2, ..., k_K >$ satisfies

$$d_{k_1} \leq d_{k_2} \leq \cdots \leq d_{k_K}. \quad (22)$$

C. Sensor sequence composition with variable transmission power

As discussed above, if the first K closest sensors achieve the required estimation error, we can compose them to a sensor sequence with the fastest convergence rate by their distances to the target. Fig. 3 illustrates an example where 6 sensors are needed to guarantee required estimation error. Suppose that their distances to the target are $d_1 \leq d_2 \leq ... \leq d_6$. The sensor sequence $< s_1, s_2, s_3, s_4, s_5, s_6 >$ is the one with the fastest convergence rate. If all the sensors in an estimation sequence transmit with the same power level and by which they can form a connected graph, then the sequence with the fastest convergence rate has the same total transmission power cost as that of any other allowable estimation sequence. In some cases, the transmission power can be dynamically adjusted and the minimum power is used to achieve the required transmission quality which is dependent on the distance between the sending sensor and the receiving sensor. A commonly used model for adjustable transmission power between sensor i and receiver j is as

$$p^t_{ij} = C \times d^{\beta}_{ij}, \quad (23)$$

where C is a constant, β the radio transmission decay exponent and d_{ij} the Euclidean distance between i and j. In such cases, we may need to compose a sensor sequence for the first K closest sensors with the least total transmission power, denoted by $\overline{K}^t$. We note that $\overline{K}^t$ may not be unique.

For simplicity, we assume that the first K closest sensors can form a connected graph $G = (V, E)$ with the edge cost $c(i, j) \equiv p^t_{ij}$ between two vertices i and j. If node i cannot reach node j with the largest power, we set $c(i, j) = \infty$. The sequence with the least total transmission power, $\overline{K}^t =< k^t_1, k^t_2, ..., k^t_K >$, is a *simple path* of length K in the graph with the minimum cost, i.e., $\overline{K}^t$ minimizes $\sum_{i=1}^{K-1} c(k^t_i, k^t_{i+1})$. Suppose that we add a virtual vertex v' connecting to all other vertices with zero costs and form a new graph $G' = (V \bigcup \{v'\}, E')$. Then our problem can be converted to a constrained version of the classic *Hamiltonian Circuit* (see e.g., [13] page 47) as to find a *hamiltonian circuit (or cycle)* starting at v' and ending at v'. The Hamiltonian Circuit problem is NP-complete and our problem can also be proven to be NP-complete with the similar arguments.

For small value of K, exhaustive search can be used to find an optimal solution: select $\overline{K}^t$ from all possible 2^K sequences. For medium to large value of K, we propose a heuristic to find a total transmission power efficient sequence as follows. Observe that the first K closest sensors are confined within a disk centered at the target with the radius of the largest sensor distance to the target. We propose to form a sequence by anti-clockwise/clockwise adding sensors. Table I describes our heuristic. We use ϕ_i to denote the angle between the line $\overline{s_i\theta}$ (connecting sensor s_i and the target) and the X-axis, $0 \leq \phi_i < 2\pi$. We use s_c to denote the last selected sensor and ϕ_c its angle. The composition begins with a trivial sequence consisting of a single arbitrary chosen sensor as the ”root sensor”. The compositions continues by anti-clockwise selecting a sensor with the smallest angle larger than ϕ_c. We use $d(s_c, s)$ to denote the Euclidean distance between sensors s_s and s, and d_{max} the maximum transmission distance when the maximum transmission power is used. If the selected sensor s is more than d_{max} away from s_c, then we come back to the ”root sensor” and the composition resumes by clockwise selecting a sensor with the smallest absolute value of the difference between its angle and ϕ_c. After the anti-clockwise and clockwise sensor selection, there may still have some sensors not included in the sequence. We then use a common insertion algorithm to insert them into the sequence. The *insertion cost* is defined as the total transmission power of the sequence when a sensor is inserted into somewhere of the sequence. The insertion operation is more computational complicated than the simple anti-clockwise/clockwise sensor selection. Given M unselected sensors and a sequence of length N, the former needs $O(MN)$ computations of insertion cost while the later needs only $O(M)$ computations of angle comparison. It is expected that most of sensors will be included in the sequence after the anti-clockwise and clockwise sensor selection if d_{max} is large enough. In the case that $d_{max} \geq 2d_K$, no insertion operation is needed, where d_K is the largest sensor distance to the target. Given the example shown in Fig. 3, the sensor sequence formed by our heuristic is as $< s_1, s_3, s_5, s_6, s_4, s_2 >$.

IV. Numerical Examples

We present some numerical examples in this section. For simplicity, we assume that all noises have the same unit variance. The unit distance is set to achieve a unit mean squared error when only one sensor is used. 100 sensors are used for estimation with their distances to the

TABLE I

HEURISTIC FOR COMPOSING A TRANSMISSION POWER EFFICIENT SENSOR SEQUENCE

select an arbitrary sensor as the "root sensor" s_r and set $\overline{K}^{\phi} =< s_r >$;
remove s_r from the available sensor set $\mathcal{S}_K$;
set $s_c = s_r$ and ϕ_c the angle of the s_r; set *status* = *anti-clockwise*;
while $\mathcal{S}_K$ is not empty,
  switch *status*,
    'anti-clockwise':
      select $s \in \mathcal{S}_K$ with the smallest angle larger than ϕ_c
      if $d(s_c, s) \le d_{max}$,
        add s to the rearmost of $\overline{K}^{\phi}$; remove s from $\mathcal{S}_K$; set $s_c = s$ and ϕ_c the angle of s;
      if $d(s_c, s) > d_{max}$,
        set *status* = *clockwise*; set $s_c = s_r$ and ϕ_c the angle of s_r;
    'clockwise':
      select $s \in \mathcal{S}_K$ such that $|\phi_c - \phi_s|$ is the minimum;
      if $d(s_c, s) \le d_{max}$,
        add s to the headmost of $\overline{K}^{\phi}$; remove s from $\mathcal{S}_K$; set $s_c = s$ and ϕ_c the angle of s;
      if $d(s_c, s) > d_{max}$,
        set *status* = *insertion*;
    'insertion':
      select $s \in \mathcal{S}_K$ with the minimum *insertion cost*;
      insert s into $\overline{K}^{\phi}$ (to the position with the minimum insertion cost); remove s from $\mathcal{S}_K$;

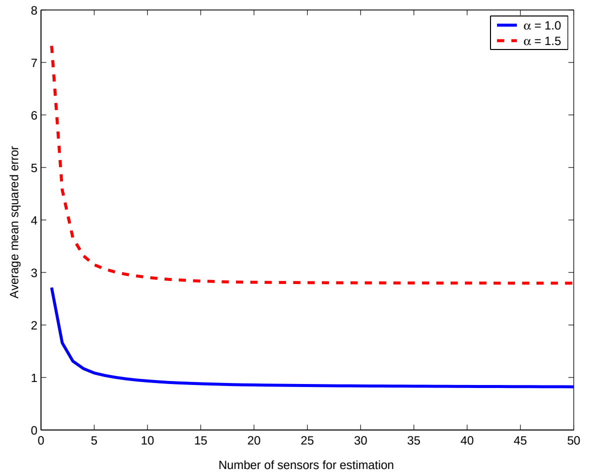

Fig. 4. Average mean squared error of the most efficient selection for a sensor sequence.

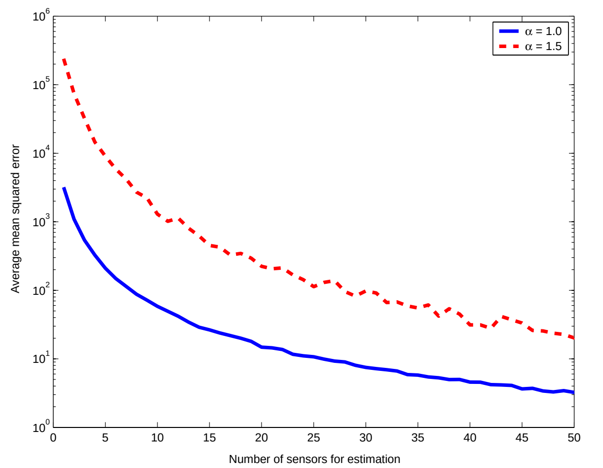

Fig. 5. Average mean squared error of the random selection for a sensor sequence.

target location uniformly random distributed from 0.1 to 100 distance units. The sensor sequence with the fastest convergence rate is to select K sensors according to Corollary 3. A random selection is to randomly select K sensors to form a sensor sequence from the 100 sensors, and the obtained MSE is averaged over 25 selections. The MSE of the two schemes are then averaged over 20 simulation runs. The results are shown in Fig. 4 and Fig. 5 for the sequence with the fastest convergence rate and the random sequence, respectively. It is first observed that the average MSE is lower for a smaller distance decay component α since low α introduce low distance attenuation. We also observe that the two kinds of sequences have similar average MSE when K is larger. However, it is not unexpected to observe that the average MSE of the sequence with the fastest convergence rate converges much faster (10 sensors) compared with that of the random selection scheme (50 sensors). Finally, the long smooth tail of the average MSE in Fig. 4 indicates that only a few number of sensors, e.g., 10 sensors in the figure, is enough to achieve a target estimation error.

We then compare power consumption for CBES and LRES. However, since we do not have exact power consumption values for sensor computations, some approximations are used in the comparison. Let a_1 denote the processing power for computing $B(k)$ and $\hat{\theta}(k)$. We first consider a simple case of using fixed and maximum transmission power. Let a_2 and a_3 denote the receive and transmit power, respectively. Though the transmit/receive power might be different for different lengths of packets, we assume that the transmit/receive power is the same in CBES and LRES. This is a reasonable assumption as the encoded $B(k)$ and $\hat{\theta}(k)$ might only be a few bytes long. When the maximum transmit power is used, the transmit

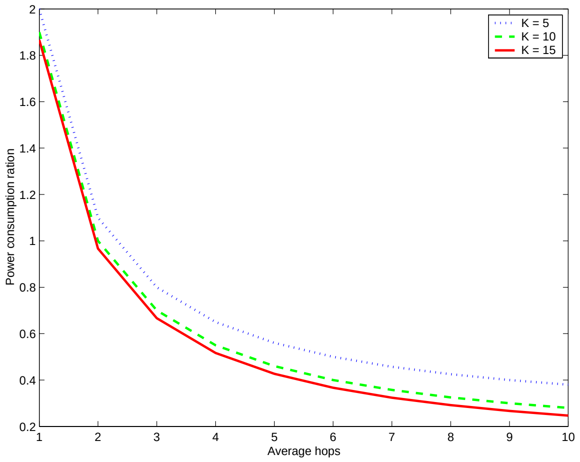

Fig. 6. Power consumption ratio with multi-hop transmissions: power consumption of LRES divided by power consumption of CBES.

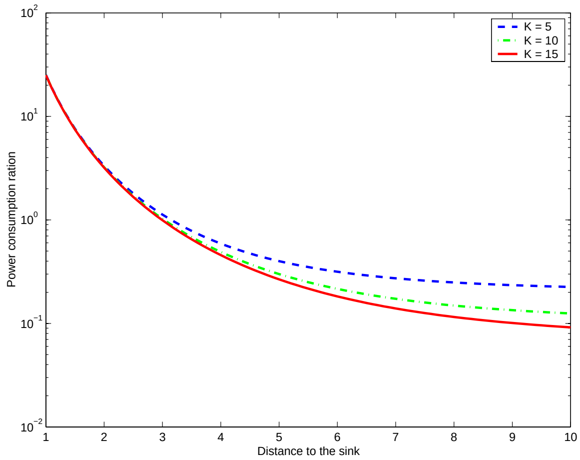

Fig. 7. Power consumption ratio with direct transmissions: power consumption of LRES divided by power consumption of CBES.

distance is also fixed to achieve a target transmission error (though it might be different for different modulation and encoding schemes). Hence we simply use *average hops* to measure the distance between local sensors to a sink (or a fusion center). Furthermore, we assume that sensors for estimation are within one hop away from each other. This is also reasonable since we only choose sensors close to the target location. Suppose there are K sensors and their average distance to the sink is i hops. Then the total power consumption for CBES is: $K \times i \times (a_2 + a_3)$; and for LRES is: $K \times (a_1 + a_2 + a_3) + i \times (a_2 + a_3)$. Some typical values of a mote sensor by Crossbow Inc., are used for computation, and they are $a_1 = 16mA$, $a_2 = 8mA$ and $a_3 = 12mA$ [14]. The power consumption ratio defined by the power consumption of LRES divided by the power consumption of CBES is plotted against the number of average hops in Fig. 6. It is observed that when the sink is only one hop away from the sensors, the CBES is more energy efficient; however, when the sink is three or more hops away, the LRES is more energy efficient.

We next consider a scenario that all sensors can transmit directly to the sink with adjustable transmission power. In this scenario, there is no receive power consumption in CBES as data relay is not needed. For simplicity, we assume that all K sensors have approximately the same distance d to the sink. We also assume that the distances between sensors are no more than d_{max} and the transmission power between sensors is simplified as d^{β}_{max}. We note that this assumption is a worst case and the transmission power may be greatly reduced if transmission power efficient sequence is used. The total power consumption in this scenario for CBES is: $K \times d^{\beta}$; and for LRES is: $K \times (a_1 + a_2 + d^{\beta}_{max}) + d^{\beta}$. Fig. 7 plots the power consumption ratio against the distance to the sink, where $d_{max} = 1$ and $\beta = 3.5$. Similar to the case of multi-hop transmissions, the CBES is more energy efficient when the sink is close while LRES performs much better for far away sink.

We finally compare the total transmission cost for three sensor sequences discussed in Section III-C. They are the *optimal sequence* with the minimum total transmission cost among all available possible sequences (denoted by OptiSeq), the *heuristic sequence* generated by our heuristic (denoted by HeurSeq) and the *convergence sequence* with the fast convergence rate (denoted by ConvSeq). We randomly scatter K sensors within a disk centered at the target with the diameter equal to the maximum transmission distance. Table II compares the average total transmission power of the three sequences. The *cputime* is the computation time measured by the Matlab *cputime* function when executing the exhaustive search for finding the optimal sequence. We note that the computation time for finding a heuristic sequence and a convergence sequence is negligible and hence we do not list them in the table. From the table, we can observe that the gap between the proposed heuristic and the optimal one is large for small values of K. However, when the value of K increases, the computation time of the optimal sequence increases exponentially and when the value of K is greater than 9, the computation time may take several hours. In such a case, we have to use a heuristic method to trade-off between computation time and the performance. On the other hand, we observe that the heuristic sequence performs much better than the convergence sequence. This is also illustrated by Fig. 8 which compares the average total transmission power for the two sequences. Since the sensors are randomly scattered, the distances between two consecutive sensors in the convergence sequence are randomly distributed also: Two consecutive sensors in the convergence sequence may be separated far away and many zigzag transmissions have to be performed. The proposed heuristic tries to restrict transmission only between two nearby sensors along with the circular direction and hence the total transmission power can be greatly reduced.

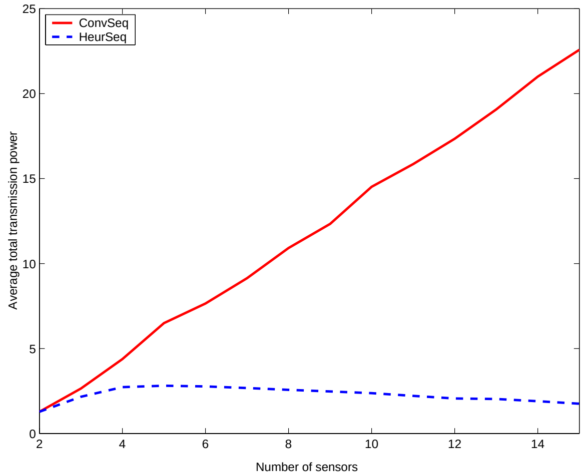

Fig. 8. Average total transmission of the two sequences.

TABLE II
COMPARISON OF AVERAGE TOTAL TRANSMISSION POWER.

	cputime	Average total transmission power		
K	OptiSeq	OptiSeq	HeurSeq	ConvSeq
3	2.0e-4	1.59	2.41	2.96
4	1.5e-3	1.30	2.53	3.71
5	8.1e-3	1.53	3.02	6.86
6	7.2e-2	1.41	2.83	7.58
7	5.9e-1	1.29	3.05	9.44
8	7.1	1.23	2.54	11.29

V. RELATED WORK AND DISCUSSIONS

Recently, the problem of how to use geographically distributed sensors to collectively accomplish estimation has re-attracted a lot of research attention in the context of wireless sensor networks [4][5][6][7] [8][15][16][17]. In all of the above work, the parameter to be estimated is considered to be consistent and does not decay with distances. The local measurement of a sensor is given by

$$x_k = \theta + n_k,\ k = 1, 2, ..., K.$$

Each sensor sends its quantized (and modulated and/or encoded) measurements to a fusion center where the estimation is performed according to a fusion function/estimator with these messages as its inputs. The main concerns in the above work are the design and analysis of the power efficient local message function, fusion function and modulation strategy.

In [4], Luo proposes an *universal decentralized estimate scheme* where the optimal message function uses only one-bit to encode a measurement and $1/2$ available sensors are allocated to estimate the first bit of the unknown parameter, $1/4$ of the available sensors are allocated to estimate the second bit of the unknown parameter and so on. In [5], Luo extends his work in [4], and proposes an *isotropic universal decentralized estimate scheme* where the local message function and the fusion function are independent on sensor index, noise distribution, network size, or network topology. Actually, the proposed scheme allows sensors to operate identically and autonomously without global network information.

In [6][7][8], Xiao and Luo propose a decentralized estimate scheme in an *inhomogeneous* sensing environment where sensors may have different local signal-to-noise ratio (SNR). In the proposed scheme, each sensor encodes its measurement to a small number of bits with length proportional to the logarithm of its local SNR. The proposed scheme is also universal in that the probability distribution functions of the noises are not required for the local message function and the final fusion function.

In [15], Krasnopeev, Xiao and Luo extend their previous work [4]– [8] and propose a decentralized estimation scheme without the assumption that the sensors' noises are independent and uncorrelated. Instead, they consider the situation that measurements are corrupted by *correlated* additive noises. The local message function uses a simple quantization strategy needing no knowledge of the noise. The covariance matrix of noises is assumed to be known in the fusion center and is included in the fusion function.

In [16][17], Xiao et al combines the decentralized estimation with the transmission power allocation. Given the universal decentralized estimation scheme [8] and an uncoded *quadrature amplitude modulation* (QAM), the optimal message function and transmission power level are determined to minimize the total transmission power while still meeting the given estimation error requirement. Sensors with bad measurements or bad channels are suggested to reduce their quantization levels or become inactive.

Obviously, we see that our estimation problem differs with the above in two aspects. One is that our measurement model considers parameter attenuation with distances and the other is that we perform localized recursive estimation and send only the estimate result to the fusion center, instead of sending all measurements. In the paper, we have assumed the transmission of real-valued intermediate estimates. This may not be the case in practical situations. We note that even in a local recursive estimation scheme the quantization, modulation and encoding, and transmission (in noisy channels) of the intermediate estimates and error variance are also indispensable to the performance of the estimator. Their impacts on the performance of a localized recursive estimator are one of our future work.

VI. CONCLUDING REMARKS

In this paper, we have proposed a localized recursive estimation scheme for parameter estimation in wireless sensor networks. Some properties of the scheme have been analyzed and the sensor sequence with the fastest convergence rate has been identified. In the case of adjustable transmission power, a heuristic has been proposed to find a sensor sequence with the minimum total transmission power when performing the recursive estimation. Numerical examples have been used to compare the performance of the proposed scheme with that of a centralized estimation scheme and have shown the effectiveness of the proposed heuristic.

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks, Elsevier Publishers*, vol. 39, no. 4, pp. 393–422, 2002.

[2] M. C. Vuran, Ö. B. Akan and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Computer Networks, Elsevier Publishers*, vol. 45, pp. 245–259, 2004.

[3] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy constrined ad hoc wireless networks," *IEEE Wireless Communications*, vol. 9, no. 4, pp. 8–27, August 2002.

[4] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. on Information Theory*, vol. 51, no. 6, pp. 2210–2219, 2005.

[5] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 735–744, 2005.

[6] Z.-Q. Luo and J.-J. Xiao, "Decentralized estimation in an inhomogeneous environment," *IEEE International Syposium on Information Theory (ISIT)*, pp. 517–517, 2004.

[7] Z.-Q. Luo and J.-J. Xiao, "Universal decentralized estimation in a bandwidth constrained sensor network," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2005, pp. 829–832.

[8] J.-J. Xiao and Z.-Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.

[9] G. Pottie and W. Kaiser, "Wireelss integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, May 2000.

[10] L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, *Fundamentals of Acoustics*, New Youk: John Wiley and Sons, Inc., 2000.

[11] X. Sheng and Y.-H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor ntworks," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 44–53, 2005.

[12] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*, Prentice Hall, Inc, 1995.

[13] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of Np-Completeness*, Freeman, New York, 1979.

[14] C. T. Inc., *MPR-Mote Processor Radio Borad, MIB-Mote Interface/Programming Board User's Manual (Document 7340-0021-05)*, Crossbow Technology, Inc., 2003.

[15] A. Krasnopeev, J.-J. Xiao and Z.-Q. Luo, "Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noises," *EURASIP Journal on Wireless Communications and Networking*, vol. 5, no. 4, pp. 473–482, 2005.

[16] J.-J. Xiao, S. Cui, Z.-Q. Luo and A. J. Goldsmith, "Joint estimation in sensor networks under energy constraint," in *IEEE Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, 2005, pp. 264–271.

[17] J.-J. Xiao, S. Cui, Z.-Q. Luo and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, 2005.

Bang Wang Dr. Bang Wang obtained his Bachelor of Engineering and Master of Engineering from the Department of Electronics and Information Engineering in Huazhong University of Science and Technology (HUST) Wuhan, China in 1996 and 2000, respectively and his PhD degree in Electrical and Computer Engineering (ECE) Department of National University of Singapore (NUS) Singapore in 2004. He is now working as a research fellow in the ECE Dept. of NUS on wireless sensor networks. His research interests include coverage issues, distributed signal processing, resource allocation and optimization algorithms in wireless networks.

Kee Chaing Chua Dr. Kee Chaing Chua received a PhD degree in Electrical Engineering from the University of Auckland, New Zealand, in 1990. Following this, he joined the Department of Electrical Engineering at the National University of Singapore (NUS) as a Lecturer, became a Senior Lecturer in 1995 and an Associate Professor in 1999. From 1995 to 2000, he was the Deputy Director of the Center for Wireless Communications (now Institute for Infocomm Research), a national telecommunication R&D institute funded by the Singapore Agency for Science, Technology and Research. From 2001 to 2003, he was on leave of absence from NUS to work at Siemens Singapore where he was the founding head of the ICM Mobile Core R&D department. Dr. Chua has carried out research in various areas of communication networks and has published more than 150 papers in these areas in international refereed journals and conferences. His current interests are in ensuring end-to-end quality of service in wireless and optical networks. He is a recipient of an IEEE 3rd Millennium medal.

Vikram Srinivasan Dr. Vikram Srinivasan obtained his Bachelor of Science in Physics from the University of Chennai in 1994 and an M.E in Electrical Communications Engineering from the Indian Institute of Science, Bangalore in 1998. He obtained his PhD in Electrical and Computer Engineering from the University of California at San Diego in 2003. He is currently an Assistant Professor in the Electrical and Computer Engineering Department at the National University of Singapore. His research interests are broadly in the area of wireless networks.