Coherent Acoustic Array Processing and Localization on Wireless Sensor Networks

JOE C. CHEN, MEMBER, IEEE, LEN YIP, JEREMY ELSON, HANBIAO WANG, MEMBER, IEEE, DANIELA MANIEZZO, STUDENT MEMBER, IEEE, RALPH E. HUDSON, KUNG YAO, FELLOW, IEEE, AND DEBORAH ESTRIN, SENIOR MEMBER, IEEE

Invited Paper

Advances in microelectronics, array processing, and wireless networking have motivated the analysis and design of low-cost integrated sensing, computing, and communicating nodes capable of performing various demanding collaborative space–time processing tasks. In this paper, we consider the problem of coherent acoustic sensor array processing and localization on distributed wireless sensor networks. We first introduce some basic concepts of beamforming and localization for wide-band acoustic sources. A review of various known localization algorithms based on time-delay followed by least-squares estimations as well as the maximum–likelihood method is given. Issues related to practical implementation of coherent array processing, including the need for fine-grain time synchronization, are discussed. Then we describe the implementation of a Linux-based wireless networked acoustic sensor array testbed, utilizing commercially available iPAQs with built-in microphones, codecs, and microprocessors, plus wireless Ethernet cards, to perform acoustic source localization. Various field-measured results using two localization algorithms show the effectiveness of the proposed testbed. An extensive list of references related to this work is also included.

Keywords*—*Ad hoc network, beamforming, distributed sensor network, microphone array, source localization, time synchronization, wireless network.

I. INTRODUCTION

Recent developments in integrated circuit (IC) technology has allowed the construction of low-cost small sensor nodes with signal processing and wireless communication capabilities that can form distributed wireless sensor network systems. These systems can be used to perform detection, localization, tracking, and identification of objects in diverse military, industrial, scientific, office, and home applications [1]–[5].

Beamforming is a space–time operation in which a waveform originating from a given source but received at spatially separated sensors are combined in a time-synchronous manner. If the propagation medium preserves sufficient coherency among the received waveforms, then the beamformed waveform can provide an enhanced signal-to-noise ratio (SNR) compared to a single sensor system. Beamforming can be used to determine the direction(s) of arrival and the location(s) of the source(s). Beamforming and localization are two interlinking problems, and many algorithms have been proposed to tackle each problem individually and jointly (i.e., localization is often needed to achieve beamforming, and some localization algorithms take the form of a beamformer).

In this paper, we consider coherent processing dealing with acoustic sources and sensors. The processing of seismic/vibrational sensor data is similar to that of acoustic sensors except for the propagation medium and unknown speed of propagation. Source types such as RF, magnetic, infrared, and visual have other distinct features and will not be considered here. Acoustic source localization and the beamforming problem is challenging due to its wide-band nature, near- and far-field geometry (relatively near/far distance of the source from the sensor array), and arbitrary array shape. In contrast, RF source can be considered to be narrow-band, far-field, and the array shape is often controlled. Many tutorial papers [6], [7] and books [8]–[10] have dealt with beamforming and localization. Recent developments in single-frame acoustic source localization can be categorized into two classes, namely, two-step methods (for single source only) with an intermediate time-delay estimation followed by a least-squares (LS) estimation [11]–[13] and

Manuscript received November 3, 2002; revised March 12, 2003. This work was supported in part by the National Science Foundation CENS program, the Intel Corporation, the Defense Advanced Research Projects Agency SensIT program AFRL Grant N374, and by NASA-Dryden Grant NCC2-374.

The authors are with the Electrical Engineering Department and Computer Science Department, University of California, Los Angeles, CA 90095 USA.

Digital Object Identifier 10.1109/JPROC.2003.814924

a single-step maximum-likelihood (ML) approach that is capable of estimating multiple source locations and directly provide beamforming outputs [14], [15]. The ML approach has an obvious advantage (at the cost of computational complexity) over the time-delay LS-type methods, since in most practical scenarios multiple signals coexist and must be considered in order to achieve high performance in localization and signal-to-interference-and-noise-ratio (SINR) enhancement.

The design of acoustic localization algorithms mainly focuses on high performance, minimal communications load, and computationally efficient and robust methods to reverberant and interference effects. In [16], a robust method for relative time-delay estimation was proposed by reformulating the problem as a linear regression of phase data and then estimating the time delay through minimization of a robust statistical error measure. When several signals coexist, the relative time delay of the dominant signal was shown to be effectively estimated using a second-order subspace method [13]. A recent application of particle filtering to acoustic source localization using a steered beamforming framework also promises efficient computations and robustness to reverberations [15]. Another attractive approach using the integration (or fusion) of distributed microphone arrays can yield high performance without demanding data transfer among nodes [17]. Unlike the aforementioned approaches that perform independent frame-to-frame estimation, a tracking framework was also developed in [18] to provide power-aware, low-latency location tracking that utilizes the historical source information (e.g., trajectory and speed) with the single-frame updates.

Besides various physical phenomena, many system constraints also limit the performance of coherent array signal processing algorithms. For instance, the system performance may suffer dramatically due to sensor location uncertainty (due to unavailable measurement in random deployment), microphone response mismatch and directivity (which may be serious for some types of microphones in some geometric configurations), and loss of signal coherence across the array (i.e., widely separated microphones may not receive the same coherent signal) [7]. In a self-organized wireless sensor network, the collected signals need to be well time-synchronized in order to yield good performance. These factors must be considered for practical implementation of the sensor network. In the past, most reported sensor network systems performing these processing operations usually involve custom-made hardware. In this paper, we propose to use iPAQ 3760s by Compaq, which are handheld, battery-powered devices normally meant to be used as personal digital assistants (PDAs). We select the iPAQ for its compactness, reasonable battery life, Linux open-source operating system support, and its availability as a commercial off-the-shelf (COTS) product. Each iPAQ has a built-in microphone and codec for sound acquisition, and it supports a spread-spectrum wireless Ethernet card for the transmission of data. As will be described later in more detail, the CPU clocks of the iPAQs are synchronized using the reference-broadcast synchronization (RBS) method [30], and the synchronized collected data is processed offline using Matlab to provide various localization and beamforming results. Extensive experiments using the wireless iPAQ testbed have been conducted to demonstrate the effectiveness of the RBS algorithm and localization/beamforming algorithms in various scenarios.

The paper is organized as follows. In Section II, we introduce the general framework of acoustic source localization and beamforming. Some practical coherent array processing issues, including fine-grained time synchronization, are considered in Section III. In Section IV, the testbed setup is described and the experimental results are demonstrated. Finally, in Section V we make some conclusions.

II. Array Signal Processing Algorithms

A. Acoustic Source Localization

For an array of R microphones simultaneously receiving M independent, spatially separated sound signals $(M < R)$, the acoustic waveform arriving at the rth microphone is given by

$$x_r(t) = \sum_{m=1}^{M} h_r^{(m)}(t) * s_m(t) + n_r(t) \qquad (1)$$

for $r = 1, \ldots, R$, where s_m is the mth source signal, $h_r^{(m)}$ is the impulse response from the mth source to the rth sensor (i.e., a delta function in free space corresponding to the time delay or a filtered response to include the reverberation effect), n_r is the additive noise, and $*$ denotes the convolution operation. For each chosen frame time (a function of the source motion and signal bandwidth), the received signal is appropriately digitized and collected into a space–time data vector $\mathbf{x} = [x_1(0), \ldots, x_1(L-1), \ldots, x_R(0), \ldots, x_R(L-1)]^T$ of length RL. The corresponding frequency spectrum data vector is then given by $\mathbf{X}(\omega_k) = [X_1(\omega_k), \ldots, X_R(\omega_k)]^T$, for $k = 0, \ldots, N-1$, where N is the number of fast Fourier transform (FFT) bins.

Denote $\boldsymbol{\Theta}$ as the estimation parameter, which in the near-field case is the source location vector $[\mathbf{r}_{s_1}^T, \ldots, \mathbf{r}_{s_M}^T]^T$ and $\mathbf{r}_{s_m}$ is the mth source location, and in the far-field case is the angle vector $[\phi_s^{(1)}, \theta_s^{(1)}, \ldots, \phi_s^{(M)}, \theta_s^{(M)}]^T$, and $\phi_s^{(m)}$ and $\theta_s^{(m)}$ are the azimuth and elevation angles of the mth source, respectively. In general, the ML estimation with additive white Gaussian noise is given by

$$\widehat{\boldsymbol{\Theta}}^{\mathrm{ML}} = \arg\max_{\boldsymbol{\Theta}} \sum_{k=1}^{\frac{N}{2}} W(k) \left\|\mathbf{P}(k, \boldsymbol{\Theta})\mathbf{X}(\omega_k)\right\|^2 \qquad (2)$$

where $W(k)$ is the weighting function by design (e.g., zero for insignificant bins and stronger weighting on dominant bins and/or high-frequency bins), $\mathbf{P}(k, \boldsymbol{\Theta}) = \mathbf{D}(k, \boldsymbol{\Theta})\mathbf{D}^\dagger(k, \boldsymbol{\Theta})$ is the projection matrix that projects the data vector into the parameter space, $\mathbf{D}^\dagger(k, \boldsymbol{\Theta}) = (\mathbf{D}(k, \boldsymbol{\Theta})^H \mathbf{D}(k, \boldsymbol{\Theta}))^{-1}\mathbf{D}(k, \boldsymbol{\Theta})^H$ is the pseudoinverse of the steering matrix $\mathbf{D}(k, \boldsymbol{\Theta})$ [14], and only the positive frequency bins are considered (negative frequencies are simply mirror images for real-valued signals). Note that no closed-form solution is available for (2), and efficient iterative computational methods have been proposed, including the alternating projection [14], particle filtering [15], and the SAGE method [20]. For the actual implementation of the approximate maximum-likelihood (AML) algorithm in the experiments shown in Section IV-B, we select the 100 highest frequency bins of the received sensor data for processing. A finite grid search [with equal vertical and horizontal spacing of 0.001 m for source localization and 0.3° for direction-of-arrival (DOA) estimation] is then conducted to find the global maximum of the log-likelihood function. Note that a refinement of the estimation can be performed using iterative methods, but is not considered in this paper.

In the case of a single source, the time-delay-LS methods can also be applied [11]–[13]. The time difference of arrival (TDOA) can be estimated by using various correlation operations among sensors [16] or a blind beamforming method proposed in [13]. Without a loss of generality, we choose $r = 1$ as the reference sensor for differential time delays. Let the reference sensor be the origin of the coordinate system for simplicity. The TDOA for R sensors satisfies

$$t_{r1} = t_r - t_1 = \frac{\|\mathbf{r}_s - \mathbf{r}_r\| - \|\mathbf{r}_s - \mathbf{r}_1\|}{v} \tag{3}$$

for $r = 2, \ldots, R$, where $\mathbf{r}_s = [x_s, y_s, z_s]^T$ and $\mathbf{r}_r = [x_r, y_r, z_r]^T$ are the source location and rth sensor location, respectively, and v is the speed of propagation. This is a set of $R - 1$ nonlinear equations, which after some manipulations can be formulated as a closed-form LS solution (with or without constraints) [11]–[13] of the form $\mathbf{A}\mathbf{y} = \mathbf{b}$, where the system matrix $\mathbf{A}$ contains the sensor locations and measured TDOAs, the unknown vector $\mathbf{y}$ contains the source location, source range, and v (if unknown), and the vector $\mathbf{b}$ is a function of sensor locations. An overdetermined LS solution can be given in the case of six or more sensors (for three-dimensional localization and unknown v). For a two-dimensional scenario, the constrained LS (CLS) solution to estimate the source location has the form of

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_2^T & t_{21} \\ \mathbf{r}_3^T & t_{31} \\ \vdots & \vdots \\ \mathbf{r}_R^T & t_{R1} \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{r}_s \\ v\|\mathbf{r}_s\| \end{bmatrix}$$

$$\mathbf{b} = (1/2) \begin{bmatrix} (vt_{21})^2 + \|\mathbf{r}_2\|^2 \\ (vt_{31})^2 + \|\mathbf{r}_3\|^2 \\ \vdots \\ (vt_{R1})^2 + \|\mathbf{r}_R\|^2 \end{bmatrix}.$$

The constraint of the previous equations is given by $v^2\|\mathbf{B}\mathbf{y}\| = \|\mathbf{f}^T\mathbf{y}\|$, where

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{f} = [0\ 0\ 1]^T.$$

The CLS solution can be obtained by the Lagrangian multiplier method. The solution of the unknown vector $\mathbf{y}$ can be given by $\mathbf{y} = (\mathbf{A}^T\mathbf{A} + \lambda v^2\mathbf{B} - \lambda\mathbf{f}\mathbf{f}^T)^{-1}\mathbf{A}^T\mathbf{b}$. The Lagrangian multiplier λ can be obtained by substituting $\mathbf{y}$ back to the previous constraint equation. The resulting equation is a fourth-order equation. We then select the root which gives estimate the most physical meaning. A similar LS formulation is also available for DOA estimation [19].

B. Wide-Band Beamforming

The main purpose of beamforming is to improve the SINR, which is often performed after a desired source location is obtained (except for the blind beamforming methods). In the most general sense of digital wide-band beamforming in the time-domain, the digitized received array signal is combined with appropriate delays and weighting to form the beamformer output

$$y(n) = \sum_{r=1}^{R} \sum_{\ell=0}^{L-1} w_{r\ell} x_r(n - \ell), \tag{4}$$

where $w_{r\ell}$ is the chosen beamforming weight to satisfy some criterion, and x_r here denotes the digitized version of the received signal. Numerous criteria exist in the design of the beamforming weight, including maximum SINR with frequency and spatial constraints. Other robust blind beamforming methods have also been proposed to enhance SINR without the knowledge of the sensor responses and locations. For instance, the blind maximum power (MP) beamformer in [13] obtains array weights from the dominant eigenvector (or singular vector) associated with the largest eigenvalue (or singular value) of the space–time sample correlation (or data) matrix. This approach not only collects the maximum power of the dominant source, but also provides some rejection of other interferences and noise.

In some cases, especially for multiple sources, frequency-domain beamforming may be more attractive for acoustic signals due to their wide-band nature. This is especially advantageous when a ML localization algorithm is used *a priori*, since the beamforming output is a direct result of the ML source signal vector estimate $\widehat{\mathbf{S}}^{\mathrm{ML}}(\omega_k)$ given by

$$\mathbf{Y}(\omega_k) = \widehat{\mathbf{S}}^{\mathrm{ML}}(\omega_k) = \mathbf{D}^{\dagger}\left(k, \widehat{\mathbf{\Theta}}^{\mathrm{ML}}\right)\mathbf{X}(\omega_k) \tag{5}$$

where $\mathbf{Y}(\omega_k)$ is the beamformed spectrum vector for the M sources [14]. The ML beamformer in effect performs signal separation by utilizing the physical separation of the sources, and for each source signal, the SINR is maximized in the ML sense. When only a single source exists, $\mathbf{D}^{\dagger}(k, \widehat{\mathbf{\Theta}}^{\mathrm{ML}})$ degenerates to a vector and only the SNR is maximized.

III. Practical Coherent Array Processing Issues

Practical coherent array processing design must consider ill effects of the propagation medium and channel disturbances and the imperfections of the sensor array system. ML array processing under reverberation and various types of non-Gaussian channel noises have been considered in

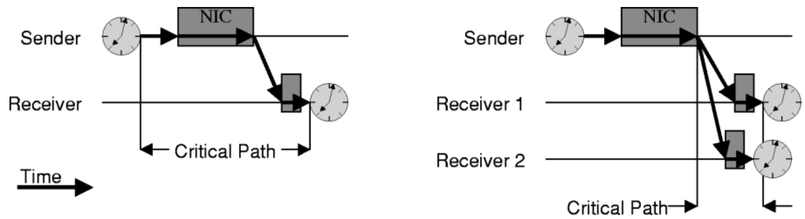


Fig. 1. A critical path analysis for traditional time synchronization protocols (*left*) and RBS (*right*). For traditional protocols working on a LAN, the largest contributions to nondeterministic latency are the send time (from the sender's clock read to delivery of the packet to its network interface card (NIC), including protocol processing) and access time (the delay in the NIC until the channel becomes free). The receive time tends to be much smaller than the send time because the clock can be read at interrupt time, before protocol processing. In the RBS, the critical path length is shortened to include only the time from the injection of the packet into the channel to the last clock read.

[21]–[23], under imperfect spatial coherence across the array [24], and uncalibrated/partially calibrated sensor array [25], [26]. Equalization of the mismatched sensor responses under known conditions [27] and blind conditions [28] using various blind equalization methods proposed for wireless communication systems [29] can also be used here.

Another important issue for coherent array processing is the need for fine-grain time synchronization among the sensors. Time synchronization comes easily for a sensor array connected by wires to a multichannel A/D converter. However, precise time synchronization becomes a critical performance limiting factor when low-cost sensor nodes digitize their signals locally and are used for coherent processing in a digital wireless network. Over the years, many protocols have been designed for maintaining synchronization of physical clocks over computer networks. Most share the same basic design: a server periodically sends a message containing its current clock value to a client. A simple one-way message suffices if the typical latency from server to client is small compared to the desired accuracy. A common extension is to use a client request followed by a server's response. By measuring the total round-trip time of the two packets, the client can estimate the one-way latency. This allows for more accurate synchronization by accounting for the time that elapses between the server's creation of a timestamp and the client's reception of it. The network time protocol (NTP) [31] is a ubiquitously adopted protocol for Internet time synchronization that exemplifies this design.

Synchronization algorithms that measure round-trip delay have one major weakness: their one-way latency estimate is confounded by differences in the forward and reverse path delays. Of course, this is very likely in the Internet, where time-varying cross traffic causes differences in queueing delay from message to message. However, such effects are also seen in LANs, where jitter on the order of tens to thousands of microseconds can be introduced by the medium-access control (MAC) layer. The value of nondeterministic delay varies depending on the specific type of network in use. For example, contention-based MACs (e.g., Ethernet) must wait for the channel to be clear before transmitting, and retransmit in the case of a collision. Wireless request to send/clear to send (RTS/CTS) schemes such as those in 802.11 networks require an exchange of control packets before data can be transmitted. Time-division multiple access channels require the sender to wait for its slot before transmitting. The jitter introduced is significant compared to our order-μs synchronization precision goal.

The RBS scheme [30] is based on the observation that jitter introduced within a LAN is dominated by these MAC delays, unlike the Internet, where propagation delay dominates. RBS, therefore, takes a different approach to reducing error: nodes periodically send a message to their neighbors using the network's physical-layer broadcast. Recipients use the message's arrival time as a point of reference for comparing their clocks. The message contains no explicit timestamp, nor is it important exactly when it is sent. RBS, therefore, does not synchronize a sender with a receiver, but rather synchronizes *a set of receivers with one another*. By using only receiver-to-receiver relations, the largest sources of nondeterministic latency are removed from the critical path, as seen in Fig. 1. While the delay incurred might vary unpredictably from message to message, the nature of a broadcast dictates that for a *particular* message, these quantities are *the same for all receivers*. In addition, because the residual error is often a well-behaved distribution (e.g., Gaussian), multiple reference broadcasts can be sent over time, allowing both improved precision of the phase offset estimate and correction for clock skew.

IV. Testbed Setup and Experimental Results

A. Wireless iPAQ Testbed

We selected the Compaq iPAQ 3760 Pocket PC as the testbed node for the following reasons. It has a built-in microphone and audio codec that supports sampling at 8 to 48 kHz in signed 16-b integer. Its 206-MHz StrongARM-1110 CPU, 32-MB ROM, and 64-MB RAM provide reasonable resources for signal processing. In addition, we insert an 11-Mb/s ORiNOCO PC card into each iPAQ for 802.11 wireless LAN connection. Thus, each node has integrated sensing, processing, and communication capabilities. We also chose the FAMILIAR distribution of the Linux operating system [32] for the testbed. The combination of COTS hardware and open-source operating system makes a powerful and convenient development platform.

In the testbed, fine-grained network time synchronization is realized by a daemon implementation of the RBS algo-

Table 1
Rms Error Result of the Source Localization Experiments Using AML and TDOA-CLS Methods

Result in figure no.	Configuration in figure no.	Type of source(s)	No. of frames evaluated	AML RMS Error	TDOA-CLS RMS Error
2	Not shown	vehicle	40	0.07m	0.05m
3	Not shown	music	40	0.07m	0.08m
5	4	vehicle	30	0.6m at S1, 0.3m at S4	0.7m at S1, 0.2m at S4
7	6	music	30	1.4m at S1, 0.4m at S5	1.6m at S1, 0.6m at S5
9	8	vehicle	20	0.3m	Not available
10	8	music and vehicle	20	source 1: 0.4m, source 2: 0.8m	Not available

rithm to reconcile the acoustic codecs' sample clocks with each other. The daemon simultaneously acts in both "sender" and "receiver" roles. Every 10 s (slightly randomized to avoid unintended synchronization), each daemon emits a reference broadcast pulse packet with a sequence number and sender ID. The daemon also watches for such packets to arrive; it timestamps them and periodically sends a report of these timestamps back to the pulse sender along with its receiver ID. The pulse sender collects all of the pulse reception reports and computes clock conversion parameters between each pair of nodes that heard its broadcasts. These parameters are then broadcast back to local neighbors. The RBS daemons that receive these parameters make them available to users. RBS never sets the nodes' clocks, but rather provides a user library that converts UNIX `timevals` from one node ID to another.

To test the precision of the time synchronization on our platform, we connected a general purpose input/output (GPIO) output from each of two iPAQs to an external logic analyzer. The analyzer was programmed to report the time difference between two pulses seen on each of its input channels. In each trial, we used RBS clock conversion parameters to command each iPAQ to raise their GPIO lines high at the same time. We ran a total of 325 trials, each separated by about 8 s, for a total test period of about 45 min. RBS achieved a mean 1.26 ± 1.11 μs synchronization error. We believe this is primarily limited by the iPAQ's clock resolution under Linux, which is 1 μs.

All nodes are organized into clusters. The cluster head commands other nodes to collect the same number of acoustic data samples starting from the same time. However, the low-end consumer-grade audio codecs on iPAQ 3760s have large nondeterministic latencies when they are asked to start recording. Simply starting recording at the same time on all sensor nodes does not guarantee getting audio data starting from the same time even if all sensor nodes' CPUs are perfectly synchronized. We can avoid this problem by using the "audio server" [33]. The audio server is a daemon that continuously runs the audio codec for sampling. In addition, it timestamps and buffers the most recent 10 s of audio data, and makes the data available to user applications through a library function. The cluster head picks a recent local timestamp, converts it to each sensor node's local timestamp, and then sends each sensor node a data request with the specified sample numbers and the specified starting time in terms of the sensor node's local time. The sensor node simply requests from the audio server the specified number of audio data samples starting from the specified local time, then sends them back to the cluster head. In this way, the cluster head collects from all sensor nodes the same number of audio data samples starting from the same time.

Data requests and replies between cluster head and sensor nodes are realized by a client-server model. A server continuously runs on each sensor node. When the cluster head requests data, it creates one client thread targeting each sensor node. Requests specify the starting time and duration of the requested data. When the server receives a data request, it requests data from the audio server according to the specification, and then sends them back to the cluster head. All data request threads run concurrently; thus, data request and reply between the cluster head and each sensor node proceeds concurrently and independently.

B. Experimental Results

Using the wireless iPAQ testbed described previously, we conducted several experiments to demonstrate its effectiveness in localization/beamforming applications. At this stage, the testbed is not completely automated and does not perform real-time signal processing for localization/beamforming due to the following challenges. First, the testbed requires accurate sensor location recordings to yield accurate results, and these measurements need to be inputted to the software prior operation. Second, efficient processing algorithm software needs to be installed in the iPAQ nodes to generate real-time results. In the following experiments, the time-synchronized acoustic data is collected by the testbed and processed offline using Matlab for the purpose of algorithm verification. The experiments are conducted outdoors to avoid severe reverberation encountered in a room environment.

We considered two scenarios: direct localization for sources in the near field and DOA estimation for sources in the far field. When several subarrays are available to obtain independent DOA estimation of the same source, bearing crossings from the subarrays is used to obtain the location estimate. In the near-field case, we also assume the received signal is highly coherent across the array. Two types of algorithms tailored for both the near- and far-field cases are considered, namely, the approximated ML (AML)

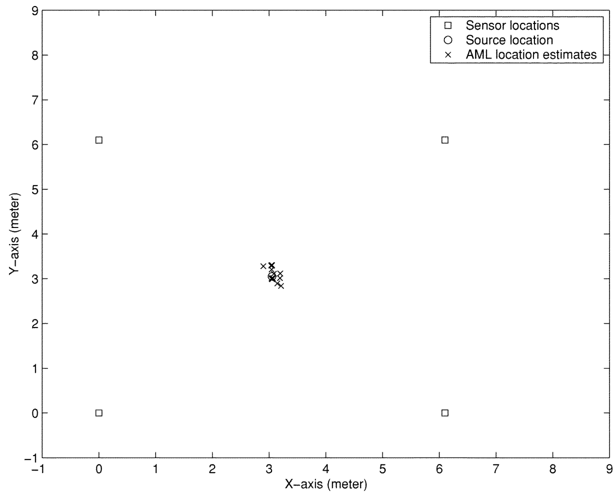


Fig. 2. AML source localization of a vehicle source.

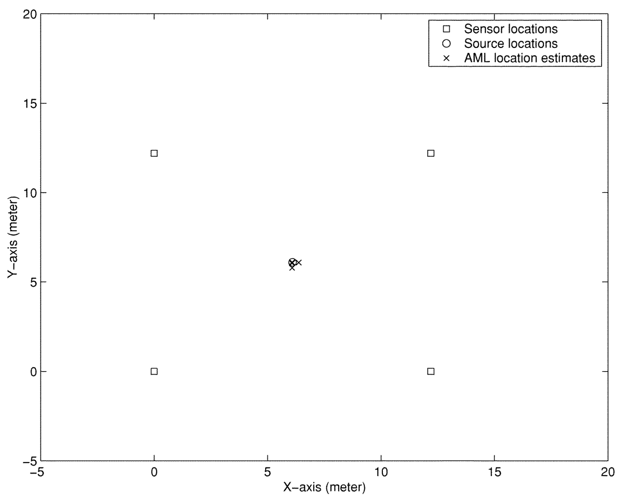


Fig. 3. AML source localization of a music source.

[14] and TDOA-CLS (constrained LS) [19]. The AML is in the form of (2) with finite frame length (limited due to possible movement of the source) and the TDOA-CLS uses the LS solution of $\mathbf{Ay} = \mathbf{b}$ with a constraint to improve the accuracy of estimation. In the case of multiple sources, only the AML method can perform estimation and the alternating projection procedure is applied. The results of the experiments are shown in Table 1 as well as the following figures.

In the first experimental setting, the source (a computer speaker) is placed in the middle of a wirelessly connected square array (each side of length L) of iPAQs. In this near-field case (relative to the array), the four nodes act as one array with intersensor spacing L of 20 ft (6.1 m). The sound of a moving light-wheeled vehicle is played through the speaker and collected by the microphone array embedded in the iPAQs. Fig. 2 shows the direct localization results of the speaker using the AML method. Then, a similar experiment is conducted using the same configuration except that L is set to be 40 ft (12.2 m). This time the loudspeaker plays prerecorded organ music, whose spectrum has a 2-kHz bandwidth with a central frequency at 1.75 kHz. The AML source localization result is shown at

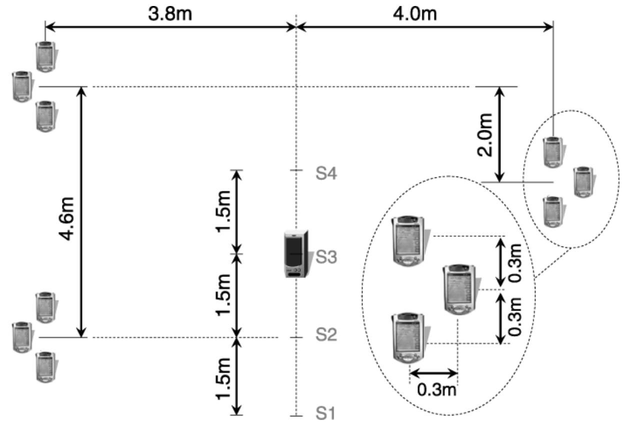


Fig. 4. Experimental setting 2: triangular subarray configuration.

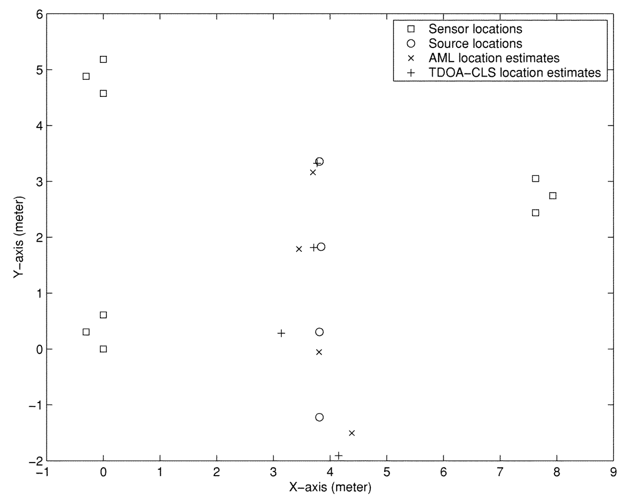


Fig. 5. Cross-bearing results of a vehicle source at different locations.

Fig. 3. From the rms errors of the previous two experiments, both algorithms show comparably promising results.

The second experimental setting is depicted in Fig. 4, where three triangular subarrays each with three iPAQs form the sensor network. In this far-field case (relative to each subarray), the DOA of the source is independently estimated in each subarray, and the bearing crossing is used to obtain the location estimate. The speaker is placed at four distinct source locations $\mathrm{S1}, \ldots, \mathrm{S4}$, simulating source movement, and the same vehicle sound is played each time. Fig. 5 depicts one snapshot (for clear illustration) of the AML and TDOA-CLS results at the four locations. We note that better results are clearly obtained when the source is inside the convex hull of the overall array. Moreover, the rms errors of Figs. 2 and 3 are much less than the rms error result here. The rms errors of direct source localization are much less because of the favorable geometry, shorter ranges, and fully coherent process in contrast to coherent DOA estimation and noncoherent cross-bearing process.

A few more far-field cases are considered. The third experimental setting is depicted in Fig. 6, where three linear subarrays each with three iPAQs form one sensor network. The speaker, this time playing the organ music sound, is placed at six distinct locations. Fig. 7 shows the one-snapshot results of

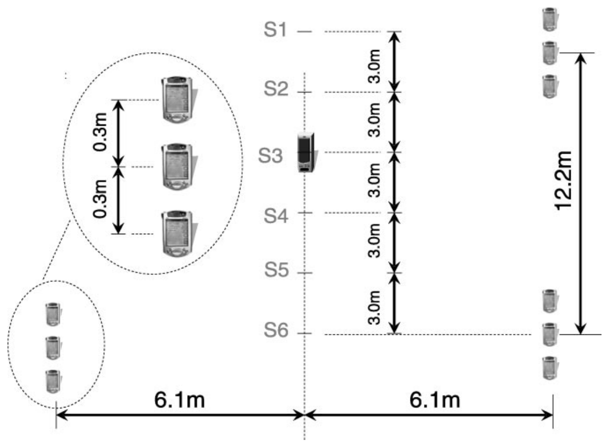


Fig. 6. Experimental setting 3: linear subarray configuration.

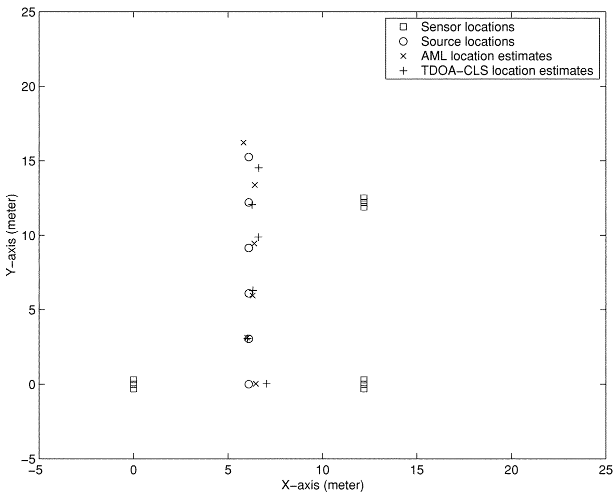


Fig. 7. Cross-bearing localization of a music source at different locations.

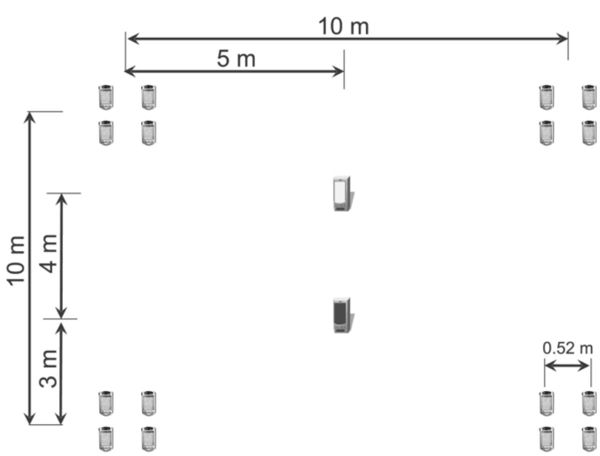


Fig. 8. Experimental setting 4: square subarray configuration.

the two algorithms at the six locations. The rms error calculation shows performance similar to the second experiment, which demonstrates that both AML and TDOA-CLS algorithm can locate different sources. The fourth experimental setting is depicted in Fig. 8, where four square subarrays each with four iPAQs form a single network. Two speakers, one

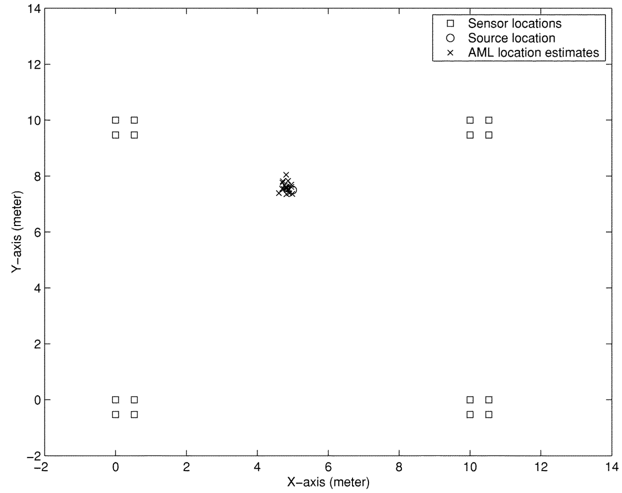


Fig. 9. AML cross-bearing localization of a vehicle source.

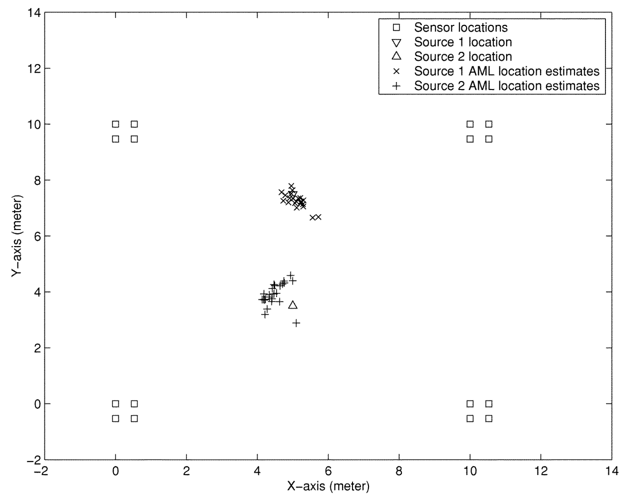


Fig. 10. AML cross-bearing localization of two sources using alternating projection.

playing the vehicle sound and the other one playing the music sound simultaneously, are placed inside the convex hull of the overall array—first, when only speaker one is playing, as shown in Fig. 9; then, when both sources are playing simultaneously, as shown in Fig. 10. Comparing Figs. 9 and 10, we see good multisource localization, but not as good as the performance with only a single source.

Note when the number of subarray element increases, the localization accuracy of the results reported above improves, which agrees with the Cramer–Rao bound analysis reported in [34].

V. CONCLUSION

In this paper, a wireless sensor network testbed is successfully implemented using COTS products. Promising localization results are shown in the offline processing of the collected data. The experiments show the effectiveness and proper joint operation of the localization/beamforming and time synchronization algorithms to yield good results. Future work includes real-time processing on the iPAQ nodes.

REFERENCES

[1] J. Agre and L. Clare, "An integrated architecture for cooperative sensing and networks," *Computer*, vol. 33, pp. 106–108, May 2000.

[2] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Commun. ACM*, vol. 43, pp. 51–58, May 2000.

[3] "Special issue on collaborative signal and information processing in microsensor networks," *IEEE Signal Processing Mag.*, vol. 19, pp. 13–85, Mar. 2002.

[4] "Special issue on advances in information technology for high performance and computational intensive distributed sensor networks," *Int. J. High Perform. Comput. Applicat.*, vol. 16, pp. 203–353, Fall 2002.

[5] "Special issue on sensor networks," *Eur. J. Appl. Signal Process.*, Mar. 2003.

[6] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, pp. 4–24, Apr. 1988.

[7] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Processing Mag.*, vol. 19, pp. 30–39, Mar. 2002.

[8] D. H. Johnson and D. E. Judgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[9] M. S. Brandstein and D. B. Ward, *Microphone Arrays: Techniques and Applications*. Berlin, Germany: Springer-Verlag, 2001.

[10] H. L. Van Trees, *Optimum Array Processing*. New York: Wiley, 2002.

[11] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1661–1669, Dec. 1987.

[12] M. S. Brandstein, J. E. Adcock, and H. F. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech Audio Processing*, vol. 5, pp. 45–50, Jan. 1997.

[13] K. Yao, R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1555–1567, Oct. 1998.

[14] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. Signal Processing*, vol. 50, pp. 1843–1854, Aug. 2002.

[15] D. B. Ward and R. C. Williamson, "Particle filter beamforming for acoustic source localization in a reverberant environment," in *Proc. IEEE ICASSP*, vol. 2, 2002, pp. 1777–1780.

[16] M. S. Brandstein and H. F. Silverman, "A robust method for speech signal time-delay estimation in reverberant rooms," in *Proc. IEEE ICASSP*, vol. 1, 1997, pp. 375–378.

[17] P. Aarabi, "Robust sound localization using distributed microphone arrays," in *Proc. ISIF*, vol. WeB3, 2001, pp. 3–8.

[18] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Processing Mag.*, vol. 19, pp. 61–72, Mar. 2002.

[19] H. Wang, L. Yip, D. Maniezzo, J. C. Chen, R. E. Hudson, J. Elson, and K. Yao, "A wireless time-synchronized COTS sensor platform part II: applications to beamforming," in *Proc. IEEE CAS Workshop Wireless Communications and Networking*, Sept. 2002.

[20] P. J. Chung and J. F. Böhme, "Comparative convergence analysis of EM and SAGE algorithms in DOA estimation," *IEEE Trans. Signal Processing*, vol. 49, pp. 2940–2949, Nov. 2001.

[21] M. Cdervall and R. L. Moses, "Efficient maximum likelihood DOA estimation for signal with known waveforms in the presence of multipath," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 45, pp. 808–811, Mar. 1997.

[22] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-Gaussian noise and Gaussian mixtures," *IEEE Trans. Signal Processing*, vol. 45, pp. 3520–3535, Dec. 2000.

[23] P. G. Georgiou and C. Kyriakakis, "A novel model for reverberant signals: Robust maximum likelihood localization of real signals based on a sub-Gaussian model," in *Conf. Rec. 36th Asilomar Conf.*, vol. 2, 2002, pp. 1278–1282.

[24] R. J. Kozick and B. M. Sadler, "Near-field localization of acoustic sources with imperfect coherence, distributed processing, and low communication bandwidth," *Proc. SPIE*, vol. 4393, pp. 52–63, Aug. 2001.

[25] A. J. Weiss, B. Friedlander, and D. D. Feldman, "Analysis of signal estimation using uncalibrated arrays," in *Proc. ICASSP*, vol. 3, 1995, pp. 1888–1891.

[26] A. Swindlehurst, "Optimum direction finding with partially calibrated arrays," in *Proc. ICASSP*, vol. 3, May 1995, pp. 1880–1883.

[27] A. Wang, K. Yao, R. E. Hudson, D. Korompis, F. Lorenzelli, S. F. Soli, and S. Gao, "A high performance microphone array system for hearing aid applications," in *Proc. ICASSP*, vol. 6, May 1996, pp. 3197–3200.

[28] A. L. Swindlehurst and J. H. Gunther, "Methods for blind equalization and resolution of overlapping echoes of unknown shape," *IEEE Trans. Signal Processing*, vol. 47, pp. 1245–1254, May 1999.

[29] Z. Ding and Y. Li, *Blind Equalization and Identification*. New York: Marcel Dekker, 2001.

[30] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," presented at the 5th Symp. Operating Systems Design and Implementation (OSDI 2002), Boston, MA, 2002.

[31] D. Mills, "Precision synchronization of computer network clocks," *ACM Comput. Comm. Rev.*, vol. 24, no. 2, pp. 28–43, Apr. 1994.

[32] The Familiar Project.. [Online] Available: http://familiar.handhelds.org

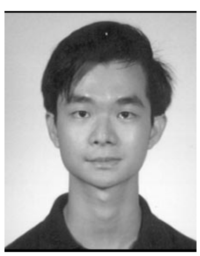
[33] L. Girod, V. Bychkobskiy, J. Elson, and D. Estrin, "Locating tiny sensors in time and space: A case study," in *Proc. Int. Conf. Computer Design (ICCD)*, 2002, pp. 214–219.

[34] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer–Rao bound," *IEEE Trans. Acoust, Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.

Joe C. Chen (Member, IEEE) was born in Taipei, Taiwan, R.O.C., in 1975. He received the B.S. (with honors), M.S., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), in 1997, 1998, and 2002, respectively.

From 1997 to 2002, he was with the sensors and electronics systems group of Raytheon Systems Company (formerly Hughes Aircraft), El Segundo, CA. From 1998 to 2002, he was a Research Assistant at UCLA, and from 2001 to 2002, he was a Teaching Assistant at UCLA. In 2002, he joined Northrop Grumman Space Technology (formerly TRW), Redondo Beach, CA, as a Senior Member of the technical staff. His research interests include estimation theory and statistical signal processing as applied to sensor array systems, communication systems, and radar.

Dr. Chen is a member of Tao Beta Pi and Eta Kappa Nu.



Len Yip received the B.S. and M.S. degrees in physics from Zhongshan University, Guangzhou, China, in 1993 and 1996, respectively, and the M.S. degree in electrical engineering from the University of California, Los Angeles (UCLA), in 2000. He is currently working toward the Ph.D degree in electrical engineering at UCLA.

His research interests include array signal processing, statistical signal processing, and sensor networks.

Jeremy Elson received the B.S. degree from Johns Hopkins University, Baltimore, MD, in 1996 and the M.S. degree from the University of Southern California, Los Angeles, in 2000. He is currently working toward the Ph.D. degree in the computer science department, and is a member of the Center for Embedded Network Sensing (CENS), at the University of California, Los Angeles. His dissertation is on time synchronization in low-power wireless sensor networks.

His other research interests include operating system issues and programming models in reactive, self-organizing systems such as sensor networks.

Hanbiao Wang (Member, IEEE) received the B.S. degree in geophysics from the University of Science and Technology of China (USTC), Hefei, China, in 1994, the M.S. degree in geophysics and space physics, and the M.S. degree in computer science, from the University of California, Los Angeles (UCLA), in 2000 and 2002, respectively. He is currently working toward the Ph.D. degree in the Computer Science Department at UCLA, focusing on collaborative information and signal processing in sensor networks.

His research interests include designing energy and bandwidth-efficient sensor networks by intertwining tasks of networking and information processing.

Mr. Wang is a member of the Association for Computing Machinery.

Daniela Maniezzo (Student Member, IEEE) received the Laurea degree (with honors) in electronic engineering from the University of Ferrara, Ferrara, Italy, in 2000. She is currently a Ph.D. student in telecommunication engineering at the University of Ferrara.

She was involved in the design, development and evaluation of a complex environment event driven simulation of an ad hoc mobile network (SAM). In 2002, she joined the research group of Prof. Kung Yao at the Department of Electrical Engineering at the University of California, Los Angeles (UCLA), where she was involved in a NASA project for the design and performance evaluation of an unmanned aircraft telemetry system. She also worked on the design and development of a wireless audio localization system in the National Science Foundation Center for Embedded Networked Sensing (CENS). She is currently a Postdoctoral Researcher in the Computer Science Department, UCLA, and is conducting research on ad hoc networks with Prof. M. Gerla. Her current research interests include the design of MAC and routing protocols for wireless ad hoc and sensor networks as well as power control, radio channel modeling, and propagation environment problems.

Ralph E. Hudson received the B.S. degree in electrical engineering from the University of California, Berkeley, in 1960 and the Ph.D. degree from the U.S. Naval Postgraduate School, Monterey, CA, in 1969.

In the U.S. Navy, he attained the rank of Lieutenant Commander and served with the Office of Naval Research and the Naval Air Systems Command. From 1973 to 1993, he was with Hughes Aircraft Company, Los Angeles, CA; since then, he has been a Research Associate in the Electrical Engineering Department of the University of California, Los Angeles. His research interests include signal and acoustic and seismic array processing, wireless radio, and radar systems.

Dr. Hudson received the Legion of Merit and Air Medal and the Hyland Patent Award in 1992.

Kung Yao (Fellow, IEEE) received the B.S.E., M.A., and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1961, 1963, and 1965, respectively.

He has worked at the Princeton-Penn Accelerator, the Brookhaven National Lab, and the Bell Telephone Labs, Murray Hill, NJ. He was a NAS-NRC Postdoctoral Research Fellow at the University of California, Berkeley. He was a Visiting Assistant Professor at the Massachusetts Institute of Technology, Cambridge, and a Visiting Associate Professor at Eindhoven Technical University, Eindhoven, The Netherlands. From 1985 to 1988, he was an Assistant Dean of the School of Engineering and Applied Science at the University of California, Los Angeles (UCLA). He is currently a Professor in the Electrical Engineering Department and a Member of the National Science Foundation Center for Embedded Networked Sensing (CENS) at UCLA. He has published more than 250 papers, and is the coeditor of *High Performance VLSI Signal Processing* (New York: IEEE–Wiley, 1997). His research interests include sensor array systems, digital communication theory and systems, wireless radio systems, chaos communications and system theory, and digital and array signal processing.

Dr. Yao received the IEEE Signal Processing Society's 1993 Senior Award in VLSI Signal Processing. He was on the IEEE Information Theory Society's Board of Governors and is a member of the Signal Processing System Technical Committee of the IEEE Signal Processing Society. He has been on the editorial boards of various IEEE Transactions, with the most recent being IEEE COMMUNICATIONS LETTERS.

Deborah Estrin (Senior Member, IEEE) received the B.S. degree in electrical engineering and computer science from the University of California, Berkeley, in 1980, the M.S. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1982, and the Ph.D. degree in computer science from MIT in 1985.

From 1986 to 2000, she was a Member of the Computer Science Department, University of Southern California (USC), Los Angeles. She is currently a Professor of Computer Science at the University of California, Los Angeles (UCLA), and is Director of the National Science Foundation (NSF)-funded Center for Embedded Networked Sensing (CENS). She is also a Member of the Computer Networks Division at the USC Information Sciences Institute (ISI). Her previous research interests include the design of network and routing protocols for very large, global, networks, such as scalable multicast routing and transport protocols, self-configuring protocol mechanisms for scalability and robustness, and tools and methods for designing and studying large scale networks. More recently, she has been collaborating with her colleagues and students at UCLA and USC/ISI to develop protocols and systems architectures needed to realize rapidly deployable and robustly operating networks of many thousands of physically embedded devices, e.g., sensor networks, toasternet, etc. She is particularly interested in the application of spatially and temporally dense embedded sensors to environmental monitoring. Professor Estrin is a Co-Principal Investigator (PI) on the Defense Advanced Research Projects Agency (DARPA)-funded SCADDS, SAMAN, and SCAN projects, and the NSF-funded SCOWR project. She was previously PI on the Virtual Internet Testbed, NSF Routing Arbiter, and RSVP-II projects at ISI. She has been an active participant in the Inter-Domain Multicast Routing WG and End-to-end research group and a Member and Study-chair for DARPA's ISAT advisory board. She is currently chairing an NRC study on Networked Embedded Computing.

Prof. Estrin is a Fellow in the Association for Computing Machinery and the American Association for the Advancement of Science. She has served on several panels for the NSF, National Academy of Sciences/NRC, DARPA, and the Office of Technology Assessment, and as a Program Committee Member for many networking-related conferences, including Sigcomm and Infocom. In 1987, she received the NSF's Presidential Young Investigator Award for her research in network interconnection and security. She has served as an Editor for the IEEE/ACM TRANSACTIONS ON NETWORKING.