Model-based Data Aggregation for Structural Monitoring Employing Smart Sensors

T. Nagayama and B. F. Spencer Jr.
Department of Civil and Environmental Engineering
University of Illinois at Urbana-Champaign
Urbana, Illinois, USA.

G. A. Agha and K. A. Mechitov
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois, USA.

Abstract— **Smart sensors densely distributed over structures can provide rich information for structural monitoring using their computational and wireless communication capabilities. One key issue in such monitoring is data aggregation. The sensors are typically sampled at high frequencies, producing large amounts of data; limited network resources (e.g., battery power, storage space, bandwidth, etc.) make acquiring and processing this data quite challenging. Efficient data aggregation with data compression is needed to achieve scalable sensor networks for structural monitoring.** *Model-based data aggregation* **is proposed using both structural and network analyses. A structural analysis algorithm, the Natural Excitation Technique, motivates adaptation of correlation function estimation to smart sensor networks. The data size is reduced by a factor of 20 to 40, depending on the degree of averaging in the aggregation. This averaging also addresses the wireless communication data loss problem. The algorithm is implemented on Mica2s and experimentally validated using a scale-model building.**

Keywords- coordinated computing; data compression; distributed algorithm; smart sensors

I. INTRODUCTION

Dense arrays of smart sensors has the potential to improve Structural Health Monitoring (SHM) dramatically using onboard computational and wireless communication capabilities. These sensors provide rich information sources, which SHM algorithms can utilize to detect, locate and assess structural damage produced by severe loading events and by progressive environmental deterioration. Information from densely instrumented structures is expected to result in the deeper insight into the physical state of a structural system.

Though smart sensors have seen several demonstrative applications to full-scale structures, none have resulted in full-fledged SHM implementations. The acquired data from several smart sensors were simply forwarded to the base station for further analysis (i.e., central processing of data), severely limiting scalability of the system. In other approaches, only primitive data processing was conducted at local nodes without internode data sharing; the data in a network was not fully exploited. To date, the capabilities of smart sensor networks have not been fully exploited for SHM.

Efficient data aggregation with data compression is needed to achieve scalable sensor networks for SHM. The scalability of the sensor network has been a critical problem for wireless sensor network (WSN), resulting in data aggregation being studied intensively. While the versatility of application independent data aggregation schemes proposed so far is appealing, the schemes do not fully account for the use of the data in SHM. Model-based data aggregation supported by structural analysis is expected to reduce data communication and to facilitate system scalability without compromising data quality from the SHM perspective.

A new model-based data aggregation scheme is proposed in this study. The Natural Excitation Technique (NExT) [1],[2], which has been employed widely in SHM strategies, motivates this scheme to estimate correlation functions in a smart sensor network. A group of smart sensors process data in a coordinated manner, as opposed to each sensor node processing data individually. This algorithm is implemented on the Mica2s and experimentally verified with a scale-model building.

II. BACKGROUND

A. Structural Health Monitoring

SHM strategies measure structural response and aim to effectively detect, locate and assess damage produced by severe loading events and by progressive environmental deterioration. Structural response reflects the structural condition as well as the excitation force. By analyzing the response data, SHM strategies are expected to determine the structural condition, such as damage existence. Once damage is detected, detailed Non-Destructive Testing (NDT), repair, and/or suspension of service follow. Well-known NDT techniques include visual inspection, eddy current testing, acoustic emission, ultrasonic testing, and radiographic inspection [3]. SHM leads to efficient maintenance of structures.

Although many methods have been proposed for SHM [4]-[8], none have proven to be sufficient for full-scale application. Many of the SHM algorithms are shown to detect damage well when a sufficient number of modes are accurately measured at all the degrees-of-freedom (DOFs). Even SHM methods which do not require response measurements at all the DOFs perform

poorly with small numbers of sensors. The more sensors used, the more reliable information SHM is expected to give about structures. However, structures are usually large and have many DOFs; accurate and thorough measurements have been impractical. Sensors have limited accuracy, installation costs, including cabling, have been prohibitively expensive. For example, The cost of installing over 350 sensors on the Tsing Ma bridge in Hong Kong was reported to be more than $8 million [9],[10]. Sensor installation cost for buildings, including labor and cabling, is estimated to be about $4,000 per sensing channel [11]. Emergence of smart sensors with wireless communication capabilities, as described in the following sections, offers the possibility of SHM with densely instrumented measurement.

B. Smart Sensors

The essential difference between a standard sensor and a smart sensor is the latter's flexible information processing capability. Each sensor has an on-board microprocessor that can be used for digital signal processing, self-diagnostics, self-identification and self-adaptation functions.

Most of the smart sensors to date employ radio-frequency-based wireless communication, while different transmission media, such as acoustic, laser and infrared transmission, also have been studied [12]. Wireless communication capability as well as battery power eliminates the cabling cost, making the sensor networks potentially scalable to a large number of nodes.

Smart sensors, primarily composed of MEMS and other commercial-off-the-shelf (COTS) components, have the potential to be produced at low cost as well as with small physical size. These features enable numerous smart sensors to be densely distributed over civil infrastructure and to capture the structure's state in detail, drawing us closer to realizing the dream of ubiquitous sensing.

The first available open hardware/software research platform is the Berkeley Mote. Under the substantial support of the US Defense Advanced Research Projects Agency (DARPA), researchers at the University of California at Berkeley have developed the open platform. The main objective is to create massively distributed sensor networks, which consist of hundreds or thousands of sensor nodes; these nodes have been termed as Smart Dust or Mote. The latest versions of the Berkeley Mote platforms, commercially available from Crossbow Technology, Inc. [13], includes the Mica2, Mica2Dot, MicaZ, and Telos.

The open nature of the platforms accelerates technical advances. Researchers can follow and customize the design, and expand its sensing modality with their own sensor boards. Sensor boards including high sensitivity acceleration sensor, strain sensor, and anti-aliasing boards have been developed for SHM applications [14]-[16]. The Berkeley Mote uses an open-source operating system, TinyOS, which is available online (http://www.tinyos.net). TinyOS is a component based operating system designed for sensor network applications on resource constrained hardware platforms, such as the Berkeley Mote. More specifically, it is designed to support the concurrency intensive operations required by networked sensors with minimal hardware requirements. The autonomous characteristics of the smart sensor can be realized by developing programs for TinyOS and then running these programs on the on-board microprocessor. Because of the open nature, numerous researchers have contributed to TinyOS enhancement; users can take advantage of the wealth of previous studies.

SHM with a dense array of smart sensors is expected to be realizable by implementing SHM algorithms on the smart sensors such as the Berkeley Motes. The limited resources of smart sensors (e.g., limited bandwidth, battery power, small RAM size, low resolution analog-to-digital converter, and modest microprocessor processor) are, however, restraining factors for intensive sensing/calculation applications such as SHM. At least, thousands of 2-byte/4-byte samples are generated at each measurement location. The limited resources prohibit direct application of traditional SHM algorithms to the smart sensor networks. Appropriate data aggregation schemes to allow scalability are needed to achieve such a SHM system.

C. Model-based Data Aggregation

Data aggregation has been proposed as a way to efficiently obtain information from WSNs. For example, combining sparse data from neighboring nodes and forwarding it in a single packet reduces the number of packets for communication and results in power efficient data aggregation. Collection of data from a smart sensor network using reliable communication protocols and shortest path routing is suitable to gather small amounts of data reliably. In this study, the focus is placed on scalable data aggregation, needed to reliably collect information from a large amount of data acquired at a dense array of smart sensors, such as structural vibration time records.

Traditional centralized data aggregation strategies to simply forward all the measured data are not scalable to a dense array of smart sensors; wireless communication to the base station is the primary bottleneck. Reducing communication requirements is necessary to achieve a scalable SHM system [17].

Simple data compression without accounting for the nature of the structural system such as downsampling and averaging are inadequate for SHM. Many of the previously proposed model-based data aggregation schemes are not directly applicable to SHM because the models do not account for the physics of the problem.

Proposed model-based data aggregation employs an application-specific model of the system to be observed. In [18], predictive data models are used to supplement actual sensor readings to give approximate answers to queries made on a sensor network. These models eliminate the need to collect data from every sensor in response to each query. Model-based data acquisition thus provides an efficient means of gathering sensor data that closely approximates the state of the system. The concept of model-based data aggregation can be utilized to obtain information efficiently from the acquired data.

For SHM analyses, a multi-degree-of-freedom (MDOF) linear time-invariant (LTI) system is oftentimes adopted as a model of a structure. Data aggregation is devised with careful consideration on this MDOF LTI system.

III. MODEL ANALYSIS

The Natural Excitation Technique (NExT) [1], [2] is widely used structural analysis technique to obtain modal information from output-only measurement of structural vibration. Because the input force to civil infrastructure is usually difficult to measure or estimate, this output-only technique is well-suited for civil infrastructure. NExT estimates the correlation functions of structural response, which can be further analyzed using modal analysis methods such as Eigensystem Realization Algorithm (ERA) [19] and Stochastic Subspace Identification (SSI) [20], [21].

Correlation functions and their frequency domain representation, spectral density functions, are commonly used analysis tools with a variety of applications. Correlation functions can be exploited to detect periodicities, to measure time delay, to locate disturbing sources, and to identify propagation paths and velocities. Spectral density function applications include identification of input, output, or system properties, identification of energy and noise source, optimum linear prediction and filtering [22]. This study specifically demonstrates the utility of correlation functions in SHM.

A. Natural Exciatation Technique

For completeness, NExT is briefly reviewed here. Consider the equation of motion in (1) under the assumption that the random responses are stationary.

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the n x n mass, damping, and stiffness matrices, respectively; $\mathbf{x}(t)$ is a n x 1 displacement vector; $\mathbf{f}(t)$ is a m x 1 force vector; $\dot{\mathbf{x}}(t)$ and $\ddot{\mathbf{x}}(t)$ are the velocity and the acceleration vectors. By multiplying displacement at reference sensor and taking expected value, (1) is transformed as follows.

$$\mathbf{M}E\left[\ddot{\mathbf{x}}(t+\tau)x_{ref}(t)\right] + \mathbf{C}E\left[\dot{\mathbf{x}}(t+\tau)x_{ref}(t)\right] \\ + \mathbf{K}E\left[\mathbf{x}(t+\tau)x_{ref}(t)\right] = E\left[\mathbf{f}(t+\tau)x_{ref}(t)\right] \tag{2}$$

Because the input force and response at the reference sensor location are uncorrelated for $\tau > 0$, the right hand side is zero. The expectation between the two signals is the correlation function. Therefore, by denoting the expectation value $E\left[x(t+\tau)y(t)\right]$ as the correlation function $R_{xy}(\tau)$, (2) is rewritten as

$$\mathbf{M}\mathbf{R}_{\ddot{x}x_{ref}}(\tau) + \mathbf{C}\mathbf{R}_{\dot{x}x_{ref}}(\tau) + \mathbf{K}\mathbf{R}_{xx_{ref}}(\tau) = \mathbf{0}, \quad \tau > 0. \tag{3}$$

From [22], when $\dot{A}(t)$ and $B(t)$ are weakly stationary processes, the following relation holds.

$$\mathbf{R}_{\dot{A}B}(\tau) = \dot{\mathbf{R}}_{AB}(\tau) \tag{4}$$

A similar relation holds for higher derivatives.

$$\mathbf{R}_{A^{(m)}B}(\tau) = \mathbf{R}^{(m)}{}_{AB}(\tau) \tag{5}$$

where the superscript, m, denotes the m-th derivative. Consequently, (3) can be rewritten as

$$\mathbf{M}\ddot{\mathbf{R}}_{xx_{ref}}(\tau) + \mathbf{C}\dot{\mathbf{R}}_{xx_{ref}}(\tau) + \mathbf{K}\mathbf{R}_{xx_{ref}}(\tau) = \mathbf{0}, \quad \tau > 0. \tag{6}$$

Thus, the correlation functions for the stationary response are shown to satisfy the equation of motion for free vibration. These correlation functions can be decomposed into modal vibration components. This fact can be directly used for modal analysis such as ERA and SSI.

B. Correlation Function Estimation

Correlation functions are, in practice, estimated from finite length records. Power and cross spectral density (PSD/CSD) functions are estimated first through the following relation.

$$G_{xy}(\omega) = \frac{1}{n_d T}\sum_{i=1}^{n_d} X_i^*(\omega) Y_i(\omega) \tag{7}$$

where $G_{xy}(\omega)$ is CSD estimation between two stationary Gaussian random processes, $x(t)$ and $y(t)$. $X(\omega)$ and $Y(\omega)$ are the Fourier transform of $x(t)$ and $y(t)$. * denotes the complex conjugate. T is time length of sample records, $x_i(t)$ and $y_i(t)$. When $n_d = 1$, the estimate has a large random error. The random error is reduced by computing an ensemble of the estimates from n_d different or partially overlapped records. The normalized rms error $\varepsilon\left[\left|G_{xy}(\omega)\right|\right]$ of the spectral density function estimation is given as

$$\varepsilon\left[\left|G_{xy}(\omega)\right|\right] = \frac{1}{\left|\gamma_{xy}\right|\sqrt{n_d}} \tag{8}$$

where γ_{xy} is the coherence function between $x(t)$ and $y(t)$, showing the degree of linearity between them. Through the averaging process, the estimation error is suppressed. Averaging of 10-20 times is common practice. The estimated spectral densities are then converted to correlation functions by inverse Fourier transform.

The correlation function estimation involves data compression. The averaging reduces the data size by a factor of 10-20, depending on the amount of averaging, n_d. In many applications, only half of correlation function data length is

utilized for the subsequent analyses. Averaging and this clipping give a total reduction of 20-40.

IV. IMPLEMENTATION

A. Data Communication strategy

The estimation of a cross correlation function needs data from two sensor nodes. Measured data needs to be transmitted from one node to the other before the processing takes place. Associated data communication can be prohibitively large without careful consideration for implementation. Data communication strategy for correlation function estimation is studied, herein.

Consider a case shown in Fig. 1 where node 1 works as a reference sensor. Assume n_s nodes, including the reference node, are measuring structural response. Each node acquires data and sends to the reference node. The reference node calculates the spectral density as in (7). This procedure is repeated n_d times for averaging. After averaging, the inverse FFT is taken to calculate the correlation function. All the calculation takes place at the reference nodes. This scheme corresponds to a centralized system where all the data are collected to a single node, with that node in charge of all the computation. When the spectral density is estimated from discrete time history records of length N, data to be transmitted through the radio is $N \times n_d \times (n_s - 1)$.

In the next scheme, the data communication strategy is devised to reduce data communication, as well as to distribute computation. After the first measurement, the reference node broadcasts the time record to all the nodes. On receiving the record, each node calculates the spectral density between its own data and the received record. The spectral density is locally stored. The nodes repeat this procedure n_d times. After each measurement, the stored value is updated by taking weighted average between the stored value and the current estimate. In this way, (7) is calculated on each node. Finally the inverse FFT is applied to the spectral density locally and then the correlation function is sent back to the reference node. Because the subsequent modal analysis such as ERA uses only half of the correlation function data length, $N/2$ data is sent back to the reference node from each node. The subsequent analyses oftentimes use only a portion of the correlation function, where estimation is accurate. Therefore, the data to be sent back can be smaller than $N/2$. The total data to be transmitted in this scheme is, at most, $N \times n_d + N/2 \times (n_s - 1)$ (see Fig. 2).

As the number of nodes increases, the advantage of the second scheme, in terms of communication requirements, becomes significant. The second approach requires data transfer of $O(N \cdot (n_d + n_s))$, while the first one needs to transmit to the base station data of the size of $O(N \cdot n_d \cdot n_s)$. The distributed implementation leverages knowledge on the application model to reduce communication requirements as well as to utilize CPU and memory in a smart sensor network efficiently.
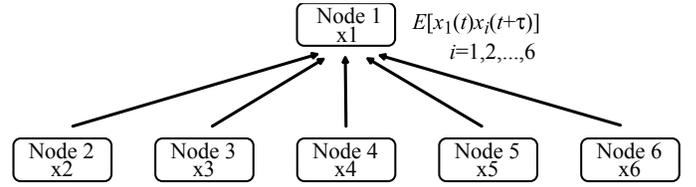

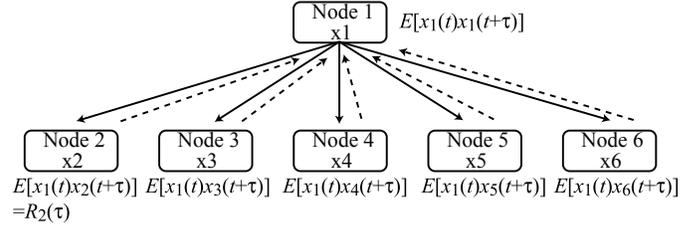Figure 1. Centralized correlation function estimation.

Figure 2. Distributed correlation function estimation.

Further consideration is necessary to accurately assess the efficacy of the two systems. Power consumption of smart sensor networks is not simply proportional to the amount of data communication. Acknowledgement messages and synchronization messages also need to be communicated. The radio listening mode consumes power even when no data is received. However, the size of the measured data is usually much larger than the size of the other messages to be sent and should be considered the primary factor in determining power consumption. Small data transfer requirements will lead to small power consumption. Note that the data communication analysis above assumes that all the nodes are in single-hop range of the reference node. This assumption is not necessarily the case for general SHM applications. However, [23] proposed a Distributed Computing Strategy (DCS) for SHM which supports this idea. In such applications, the assumption of nodes being in single hop range of a reference node is reasonable.

B. Middleware Service Requirement

Prior to implementation of the model-based data aggregation scheme on smart sensors, requirements for middleware services such as time synchronization accuracy and communication reliability need to be clarified. Appropriate middleware services which satisfy the requirements are adopted to materialize the scheme on smart sensor networks.

1) Data Loss Analysis

Data loss, which results from packet loss during wireless communication, is not assumed in the original NExT derivation [1],[2]. Some researchers have been working to have reliable communication without data loss, while others just ignore data loss effect on their analysis. The impact of the data loss on the NExT has not yet been clarified. Careful investigation in this problem may provide insight into how to accommodate communication with data loss, which is less demanding than communication without data loss, especially when a large amount of data is involved.

The averaging process in (7) is expected to reduce the impact of the data loss on spectral density and correlation function estimation. The observation noise in signals, which accompanies most of measurements, is usually addressed through this averaging process. After the raw data are processed through (7), the effect of data loss is expected to be similar to that of observation noise.

A computer simulation is conducted for a truss model (see Fig. 3), assuming various data loss levels to investigate the data loss effect. Smart sensors are assumed to be placed at the 13 nodes on the lower chord to measure the vertical acceleration. The vertical input excitation at the 6th lower chord node from the left is measured. The sampling frequency is set at 380 Hz, so that the Nyquist frequency is above the 4th natural frequency of the structure. After data is acquired, a certain percentage of data is randomly dropped to simulate data loss. Data loss is assumed to take place only in this step, where the largest amount of data is transferred among sensor nodes.

For the results of the data loss analysis to be better interpreted, the outcome is compared to that of computer simulation including observation noise (but without data loss). A band-limited white noise is added to each of the observed signals. The rms noise level is specified as a percentage of the rms of physical response.

Fig. 4 shows representative PSD function calculated with and without noise/data loss. Instead of correlation functions, their frequency domain representations (i.e., spectral densities) are compared because the spectral densities show clear peaks and zeros that reflect the modal properties of the structure. The PSD's peaks are immune to data loss or the noise because of the large signal components; however, small components near zeros are blurred. Although further investigation is necessary for quantitative judgment, these results indicate that data loss of 0.5% and 5-10% observation noise have similar impact on the PSD estimation.

The coherence function indicates the degree of linear correlation between two variables. In this computer simulation, the input excitation is applied only at node 6. Because the response of the truss structure is linear, the coherence of any two measurement points is expected to be unity over the entire frequency range. When no data loss and no observation noise are considered, the coherence function is indeed close to one, as shown in Fig.5. The isolated drops in coherence at 100 Hz and 140 Hz correspond to zeros of the system; extremely small responses of the structure at these frequencies result in numerical error, giving small coherence function.

The loss of 0.5% data reduces the coherence function except at the system's poles (see Fig. 5). Because of the random nature of data loss occurrence and the excitation, the coherence function varies from simulation to simulation. The coherence functions shown in this figure represent twenty averages. The coherence function's deviation from unity depends on the two measurement nodes, between which the function is calculated. All the investigated coherence function plots, however, support that the loss of 0.5% data affects the coherence function in a similar way as 5-10% noise addition does.
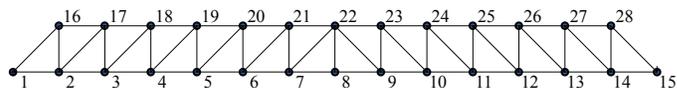

Figure 3.   Numerical simulation truss model.
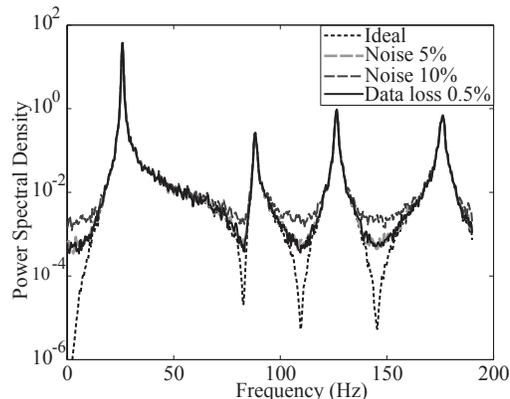

Figure 4.   Data loss effect on power spectral density function.
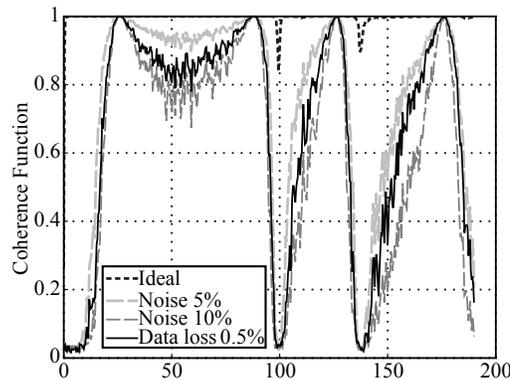

Figure 5.   Data loss effect on coherence function.

Subsequent modal and structural analyses such as identification of natural frequency and flexibility matrix estimation also indicate that a loss of 0.5% data is equivalent to 5-10% of measurement noise. Data loss is found to introduce error into the modal analysis, which also affects the subsequent analysis. Because of the averaging associated with the PSD and the CSD estimation, a certain amount of data loss can be accommodated.

Although data loss is acceptable, the position of lost data in the whole data needs to be identified. Without adequately inserting data on the receiver side in place of lost packet, data length of the whole data is not preserved. This data loss without padding seriously distorts signals and the original time history cannot be reconstructed. Consequently, data packet is sent with identifiers to indicate packet positions in the whole data so that lost data is identified in the time scale on the receiver side.

2)   Time Synchronization Analysis
Because each smart sensor has its own local clock, the measured data will have intrinsic synchronization error. The

effect of synchronization error on structural analysis is investigated in this section.

Time synchronization errors enter into correlation function estimation. Consider a cross correlation function $\mathbf{R}_{x_i x_{ref}}(\tau)$ between a response $x_i(t)$ at location i and the reference signal $x_{ref}(t)$. The sensor node at location i has a time synchronization error of t_{x_i} relative to the reference node. The correlation function, $\bar{\mathbf{R}}_{x_i x_{ref}}(\tau)$, under the synchronization error, can be written as

$$\begin{aligned}\bar{\mathbf{R}}_{x_i x_{ref}}(\tau) &= E\left[x_i\left(t - t_{x_i} + \tau\right)x_{ref}(t)\right] \\ &= \mathbf{R}_{x_i x_{ref}}\left(\tau - t_{x_i}\right)\end{aligned} \tag{9}$$

If t_{x_i} is positive, the correlation function for the interval $(0, t_{x_i})$ does not have the same characteristics as the succeeding signal. In structural analysis context, this portion has negative damping; therefore, the begging portion needs to be removed. When t_{x_i} is unknown, a segment corresponding to the maximum possible time synchronization error, $t_{x_{max}}$, is truncated from the correlation function.

The correlation functions, each having independent time synchronization errors, do not satisfy the equation of motion, (6). The correlation functions after the truncation can, however, be decomposed into modal components as

$$\mathbf{R}_{\mathbf{x}'x_{ref}'}(\tau) = \mathbf{\Phi}'\mathbf{\Lambda}\mathbf{A}$$
$$\mathbf{\Phi}' = [\varphi_1' \; \varphi_2' \; \cdots \; \varphi_{2n}']$$
$$= \begin{bmatrix} \phi_{11}\exp(-\lambda_1 t_{x_1}) & \phi_{21}\exp(-\lambda_2 t_{x_1}) & \cdots & \phi_{n1}\exp(-\lambda_n t_{x_1}) \\ \phi_{12}\exp(-\lambda_1 t_{x_2}) & \phi_{22}\exp(-\lambda_2 t_{x_2}) & \cdots & \phi_{n2}\exp(-\lambda_n t_{x_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1n}\exp(-\lambda_1 t_{x_{2n}}) & \phi_{2n}\exp(-\lambda_2 t_{x_{2n}}) & \cdots & \phi_{nn}\exp(-\lambda_n t_{x_{2n}}) \end{bmatrix}$$
$$\mathbf{\Lambda} = diag\left(\left[\exp(\lambda_1 \tau) \; \exp(\lambda_2 \tau) \; \cdots \; \exp(\lambda_{2n}\tau)\right]\right)$$
$$\lambda_i = -h_i\omega_i + j\omega_i\sqrt{1-h_i^2}$$
$$\mathbf{A} = diag\left(\left[a_1 \; a_2 \; \cdots \; a_{2n}\right]\right) \tag{10}$$

where $\mathbf{R}_{\mathbf{x}'x_{ref}'}(\tau)$ is the correlation function matrix on the interval $(t_{x_{max}}, \infty)$, ϕ_{ij} is j-th element of i-th mode shape, h_i and ω_i are i-th modal damping ratio and modal natural frequency, respectively, and a_i is a factor accounting for the relative contribution of the i-th mode in the correlation function matrix. Modal analysis techniques such as ERA identify these modal parameters. As (10) shows, the natural frequencies and damping ratios remain the same. The observed mode shapes $\mathbf{\Phi}'$ are different from the original mode shapes; changes in mode shape amplitude are negligible due to small h_i and t_{x_i}, while phase shift can be meaningful. Mode shape phases can indicate structural damage and are important modal characteristics from the SHM perspective. Time synchronization error requirements for modal analysis needs to be assessed mainly from the viewpoint of mode shape phases.

As explained in this section, the correlation function is shifted by time synchronization errors. The shifts affect mode shape estimation, while the natural frequency and damping ratio estimates are immune to the shifts. Though this analysis needs further study on estimation using finite length data, the essence of the problem is clear.

V. EXPERIMENTAL VERIFICATION

In this section, model-based data aggregation for SHM is experimentally verified using a scale-model 3-story building on a shake table (see Fig. 6). Structural response of the building under band-limited white noise ground excitation is measured by three Mica2s. Correlation functions are calculated in a distributed manner and then sent back to the base station.

A high sensitivity acceleration sensor boards, named 'Tadeo,' and a strain sensor boards are employed to record the structural response. Because the Mica2 does not have an anti-aliasing filter, anti-aliasing filter boards are used together with the 'Tadeo' and the strain sensor boards (see Fig. 7) [14]-[16]. The 'Tadeo's are installed on the second and the third floor and the strain sensor is connected to a strain gage on the first floor wall. The Mica2 with the 'Tadeo' on the third floor acts as a reference node.

Because the Mica2 has only 4 Kbytes of RAM, data to be stored on RAM is limited. Unless the memory usage is configured with virtual memory, the data record length available for processing is limited. In this experiment data length N is set to be 128. The sampling rate is 10 Hz, allowing capture of all three modes of the structural response; these natural frequencies are below 5 Hz.
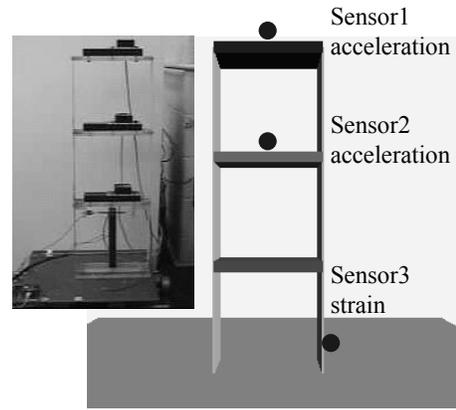

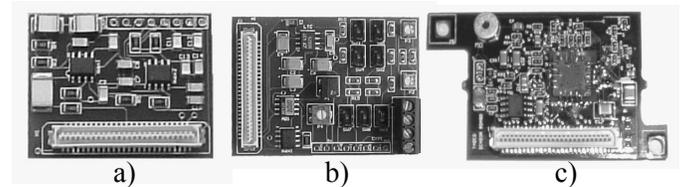Figure 6. A scale-model 3-story building.

Figure 7. Customized sensor boards. a) anti-aliasing filter board, b) strain sensor board, c) acceleration sensor board,

Reliable wireless communication is not employed for this experiment, because the averaging process in (7) accommodates data loss. Measured data were sent through the radio without repetition.

For time synchronization, the Timing-sync Protocol for Sensor Networks (TPSN) [24] module is employed under TinyOS 1.1. TPSN is reported to be able to synchronize a pair of Motes to an average accuracy of less than 20 μs and a worst-case accuracy of around 50 μs. This synchronization error is negligible for civil infrastructure, which typically has natural frequencies below 10 Hz. The maximum possible phase delay under this condition is about 0.2 degree. Also, only the first points of correlation functions are truncated, because the sampling period of implemented algorithm, 0.1 s, is much larger than the synchronization error.

Fig. 8 shows the identified correlation functions. As (6) indicates, the correlation functions seems close to free vibration signals. Because the correlation functions are not easy to judge the validity, the correlation functions are converted to spectral densities by FFT (see Fig.9). The spectral densities have three clear peaks corresponding to three natural frequencies of the building. The estimates are quite reasonable.

Finally, the correlation functions are analyzed using ERA offline. Successful modal identification indicates the validity of the correlation function estimation. The identified natural frequencies and damping ratio are close to those determined in the preliminary test using a wired data acquisition system. Furthermore, EMAC values, which are accuracy indicators for modal analysis, are reported to be close to 100% (see Table1). Because data from both acceleration and strain sensors are used in the modal analysis, the mode shapes are not easy to physically interpret. To examine the validity of the identified mode shapes, they are plotted in the complex plane. Successful modal analysis should result in mode shapes plotted almost on the real axis of the complex plane. The identified mode shapes are, indeed, almost on the real axis (see Fig. 10). The offline modal analysis of the identified correlation functions successfully detected the modal properties, which indicates that the correlation function estimation offers reliable information source to the subsequent analyses.

The data compression is significant. The raw data contained 7680 samples (=128 points x 20 averages x 3 nodes) for this experiment. The size of the processed data (i.e., the correlation functions) was 192 (=64 points x 3 nodes), giving a reduction of 40 (=7680/192). This data compression was achieved with approximately a 50% reduction in data communication. As was discussed at the end of section IV-A, while the power consumption is not a linear function of the data sent, data compression/aggregation and the associated reduction in communication generally results in power savings.

Further analysis is needed for full-fledged application. Appropriate power management, for instance, elongates the life of smart sensor networks. Functionality to sleep/wake-up, which was not implemented in this experiment, saves power. In full-scale structure applications, the responses are measured periodically, for example, once a day for 10-20 minutes.

Assuming the correlation function estimation consumes 80 mA on average for 15 minutes (i.e., 20 mAh), a AA battery with a capacity of 2000 mAh can supply power for 100 days; sleep mode power consumption is about 10 μA. With such analyses and the associated implementation, distributed correlation function estimation in this experiment will result in regularly monitoring full-scale structures using a dense array of smart sensors.
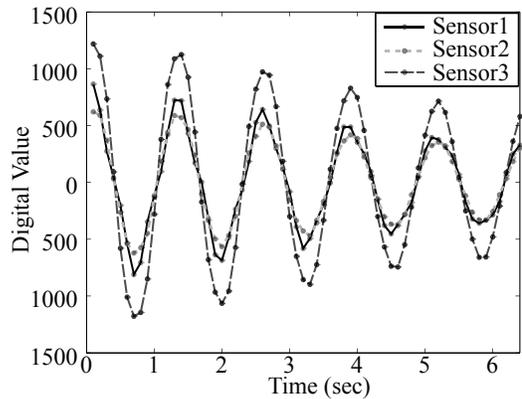

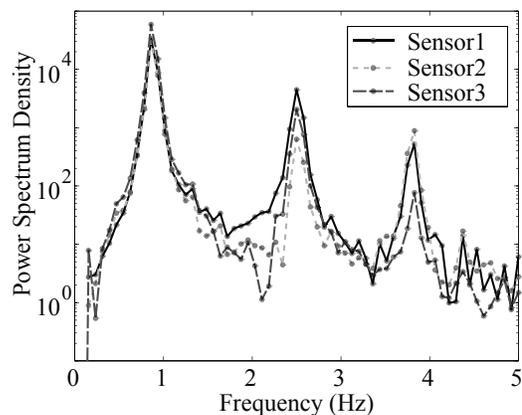Figure 8.   Estimated correlation functions.

Figure 9.   Estimated spectral density functions.
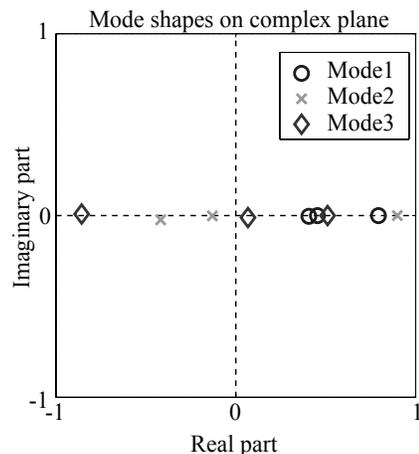

Figure 10. Identified mode shapes plotted on a complex plane.

TABLE I. IDENTIFIED MODAL PARAMETER

Mode	Natural Frequency (Hz)	Damping Ratio (%)	EMAC (%)
1st	0.7803	1.9820	100.00
2nd	2.4290	1.2508	99.98
3rd	3.7367	0.7680	99.98

VI. CONCLUSION

Model-based data aggregation supported by both structural and network analysis is proposed. This aggregation scheme utilized correlation function estimation and preserves the essential features of the data while significantly reducing data communication. Experimental verification with Mica2s mounted to a scale-model building substantiated the scheme.

The proposed data aggregation scheme is directly applicable to structural health monitoring strategies such as the recently developed DCS approach [23]. Work is currently under way to complete such implementation. Additionally, the correlation function estimation scheme can be utilized in a wide range of areas. Correlation functions and spectral densities have been used in various fields [22]. Moreover, the concept of model-based data aggregation, supported by the nature of a target system and/or objectives, has the potential to improve smart sensor networks from both the network and application specific perspectives.

REFERENCES

[1] G. H. James, T. G. Carne, and J. P. Lauffer, "The natural excitation technique for modal parameter extraction from operating wind turbine," *Report No. SAND92-1666, UC-261*, Sandia National Laboratories, Sandia, NM., 1993.

[2] G. H. James, T. G. Carne, J. P. Lauffer, and A. R. Nord, "Modal testing using natural excitation," *Proc., 10th Int. Modal Analysis Conference*, San Diego, CA, 1992.

[3] X. E. Gros, *NDT data fusion*, Arnold, London, UK , 1997.

[4] O. S. Salawu and C. Williams, "Review of full-scale dynamic testing of bridge structures." *Engineering Structures*, vol. 17, no. 2, pp. 113-121, 1995.

[5] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz, "Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review." *LA-13070-MS*, 1996.

[6] S. W. Doebling, C. R. Farrar, and M. B. Prime, "A summary review of vibration-based damage identification methods." *The Shock and Vibration Digest*, vol. 30, no. 2, pp. 91-105, 1998.

[7] S. W. Doebling and C. R. Farrar, *The state of the art in structural identification of constructed facilities*, A report by American Society of Civil Engineers Committee on Structural Identification of Constructed Facilities, 1999.

[8] H. Sohn, C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, and B. R. Nadler, "A review of structural health monitoring literature: 1996-2001" *Los Alamos National Laboratory Report, LA-13976-MS*.

[9] J. P. Lynch and K. Loh, "A summary review of wireless sensors and sensor networks for structural health monitoring," Shock and Vibration Digest, to be publised.

[10] C. R. Farrar, "Historical overview of structural health monitoring," *Lecture Notes on Structural Health Monitoring using Statistical Pattern Recognition*. Los Alamos Dynamics, Los Alamos, NM, 2001.

[11] M. Celebi, "Seismic instrumentation of buildings (with emphasis on federal buildings)" *Special GSA/USGS Project, an administrative report*, United States Geological Survey, Menlo Park, CA, 2002.

[12] S. Hollar, COTS Dust., Master's Thesis, University of California, Berkeley, CA, 2000.

[13] Crossbow Technology, Inc. http://www.xbow.com

[14] M. Ruiz-Sandoval, "*Smart sensors" for civil infrastructure systems*, Ph.D. Dissertation, University of Notre Dame, IN, 2004.

[15] T. Nagayama, M. Ruiz-Sandoval, B. F. Spencer Jr., K. A. Mechitov, G. Agha, G, "Wireless strain sensor development for civil infrastructure." *Proc., 1st Int. Workshop on Networked Sensing Systems*, Tokyo, Japan, 97-100, 2004.

[16] M. Ruiz-Sandoval, T. Nagayama, and B. F. Spencer Jr., "Sensor development using Berkeley Mote platform." *J. Earthquake Engineering*, to be published.

[17] K. Mechitov, W. Kim, G. A. Agha, and T. Nagayama, "High-frequency distributed sensing for structure monitoring." *Proc., 1st Int. Workshop on Networked Sensing Systems*, Tokyo, Japan, 101-105, 2004.

[18] A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein and W. Hong, "Model-driven data acquisition in sensor networks," *Proc., 30th Int. Conference on Very Large Data Bases*, 2004.

[19] J. N. Juang and R. S. Pappa, "An Eigensystem realization algorithm for modal parameter identification and model reduction," *J. Guidance Control and Dynamics*, vol. 8, pp. 620-627, 1985.

[20] L. Hermans and H. V. Auweraer, "Modal testing and analysis of structures under operational conditions: industrial applications." *Mechanical Systems and Signal Processing*, vol. 13, no. 2, pp. 193-216, 1999.

[21] J.-H. Yi and C.-B. Yun, "Comparative study on modal identification methods using output-only information." *Structural Engineering and Mechanics*, vol. 17, no. 3-4, pp. 445-466, 2004.

[22] J. S. Bendat and A. G. Piersol, *Random data: analysis and measurement procedures*, John Wiley and Sons, Inc. New York, NY, 2000.

[23] Y. Gao, *Structural health monitoring strategies for smart sensor networks*, Ph.D. Dissertation, University of Illinois at Urbana-Champaign, IL, 2005.

[24] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks." *Proc., 1st International Conference On Embedded Networked Sensor Systems*, Los Angeles, CA., pp. 138 – 149, 2003