Distributed Modal Identification by Regularized Auto Regressive Models

S. N. PAKZAD, G. V. ROCHA and B. YU

Abstract

Advances in Wireless sensor networks (WSN) technology have provided promising possibilities in detecting a change in the state of a structure through monitoring features of the structure that are estimated using sensor data. A set of such features are the natural vibration properties of the structure, which can be estimated from an autoregressive model (AR model) for the structure's response to ambient vibrations. Fitting an AR model to the data requires the computation of the full covariance matrix resulting in significant latency due to the low data bandwidth of WSNs, as well as high energy cost for data transmission. In this paper, a set of restrictions to the estimation of the AR model are introduced. These restrictions significantly reduce the volume of data flowing through the WSN thus reducing the latency in obtaining modal parameters and extending the battery lifetime of the WSN. Using data simulated from linear structures, the stabilization plots resulting from the restricted and full AR models are compared. The results show that the modal parameters estimated from the restricted and full models are of comparable quality. The volume of transmitted data for fitting the restricted model is significantly lower than that of the full AR model.

Introduction

With the advances in Wireless Sensor Network (WSN) technology both in terms of hardware design and software architecture, and their increasingly widespread application in structural engineering, novel data processing techniques are viewed as essential tools to enhance the performance of the integrated systems. Energy and communication bandwidth budgets are two critical issues for these networks. Efficient use of these resources is vital in installation, operation and maintenance cost and therefore viability of Wireless Sensor Networks (WSNs) for structural monitoring. The traditional paradigm for sensor networks is based on availability of abundant wired power and communication lines throughout the network. With this assumption, the cost of keeping the network powered

[1]Shamim N. Pakzad, Lehigh University, Department of Civil and Environmental Engineering, pakzad@lehigh.edu, 117 ATLSS Drive, Imbt Labs, Bethlehem, PA 18015, USA

[2]Guilherme V. Rocha, Indiana University, Department of Statistics, gvrocha@indiana.edu, 309 N. Park Av., Bloomington, IN 46217, USA

[3]Bin Yu, University of California, Berkeley, Department of Statistics, binyu@stat.berkeley.edu, 367 Evans Hall, Berkeley, CA 94720, USA

all the time is negligible, and there is little cost to transfer measured data through the network and collect it at a central processing location. Additionally in wired networks, because of the high bandwidth capacity, the latency is minimal which in turn allows for a speedy collection of the data at the central base station for processing. This paradigm is not applicable to WSNs. The power resources in a WSN are limited and expensive. The energy cost associated to communication of one bit of data is about $11,000$ times that of performing an arithmetic operation on that same bit. The communication capacity is also restricted by the radio bandwidth, which is very limited for low power networks. The wired networks have a disadvantage compared with the WSNs in that there is no computational capability in the individual sensing devices so collection of the data from the entire network at the central base station is the only available way of data processing. Distribution of analysis in-network is an optimal solution to address the limitation of both the power and bandwidth resources in a WSN.

Natural vibration properties of the structure have been the focus of many studies in structural health monitoring (SHM) and provide a great insight into the condition of the structure. A popular algorithm for obtaining the modal properties using ambient vibrations consists of autoregressive models (AR models) for the structures response [1]. The estimated autoregressive parameters contain the information that can be converted into natural frequencies, damping ratios and mode shapes. Fitting an AR model of order q to the data requires that the auto-covariance matrix (or power spectral density matrix) for the measurements be fully computed for all lags up to q. As a result, the low communication bandwidth of WSNs causes significant latency in obtaining the modal estimates.

In this paper, a set of restrictions to the estimation of the AR model are introduced to significantly reduce the volume of data owing through the WSN thus extending the lifetime of the batteries and reducing the need for data communication. Using simulations of linear vibrating systems subjected to ambient random excitations, the quality of the modal parameters estimated using the restricted model are evaluated and compared with those of the full model. More specifically, the stabilization plots of the restricted and full models are compared where the rates of convergence of the two provide insight into the performance and efficiency of the proposed method.

Multivariate Auto-Regressive Models

Multivariate autoregressive (AR) models are used to represent the evolution in time of the acceleration measured at different points of a vibrating structure. Such models are known to yield stable, reliable and accurate estimates of the dynamic properties of a structure [2, 3, 4]. The formulation of these models is briefly reviewed in this section.

If $\mathbf{u}(k)$ is the p-dimensional vector of displacements of the structure at time step k, the q-th order Auto Regressive – AR(q) – model for the acceleration ($\ddot{\mathbf{u}}(k)$) response is:

$$\ddot{\mathbf{u}}\left(k\right) \quad = \quad \sum_{j=1}^{q} \mathbf{L}_j^{(q)} \ddot{\mathbf{u}}\left(k-j\right) + \upsilon\left(k\right), \tag{1}$$

where:

$$\ddot{\mathbf{u}}(k) = \begin{bmatrix} \ddot{u}_1(k\Delta t) \\ \ddot{u}_2(k\Delta t) \\ \vdots \\ \ddot{u}_p(k\Delta t) \end{bmatrix}, \ \mathbf{L}_j^{(q)} = \begin{bmatrix} \beta_{1,1,j} & \beta_{1,2,j} & \cdots & \beta_{1,p,j} \\ \beta_{2,1,j} & \beta_{2,2,j} & \cdots & \beta_{2,p,j} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p,1,j} & \beta_{p,2,j} & \cdots & \beta_{p,p,j} \end{bmatrix}, \ \upsilon(k) = \begin{bmatrix} \upsilon_1(k\Delta t) \\ \upsilon_2(k\Delta t) \\ \vdots \\ \upsilon_p(k\Delta t) \end{bmatrix}.$$

In this equation, $\mathbf{L}_j^{(q)}$ is the matrix of coefficients for lag j of the AR(q) model, and $\upsilon(k)$ is the random ambient excitations at time $k\Delta t$.

Regularized/Constrained Auto-Regressive models

The main contribution of this paper is to propose a restricted version of the AR(q) model with the aim of reducing the communication burden on the wireless sensor network collecting the data. In addition to the reduction of traffic on the sensor network, the restricted model estimates can also result in improved quality of the estimated modal parameters by statistical regularization.

Statistical regularization

Statistical regularization consists of imposing suitable restrictions to the parameters of a model being fit to measured data. A larger number of parameters translates into more freedom to adjust the data but also into higher sensitivity to noise [5].

Letting θ denote the ideal value of a parameter representing a process and $\hat{\theta}(\mathcal{U})$ denote the estimate obtained when data set $\mathcal{U}$ is used, statisticians often use the mean squared error (MSE) as a measurement of performance of a fitting procedure:

$$MSE(\hat{\theta}) := \mathbb{E}_{\mathcal{U}}\left[\left(\hat{\theta}(\mathcal{U}) - \theta\right)^2\right] = \left[\theta - \mathbb{E}_{\mathcal{U}}\hat{\theta}(\mathcal{U})\right]^2 + \mathrm{var}_{\mathcal{U}}\left[\hat{\theta}(\mathcal{U})\right], \tag{2}$$

where the expected values and variances are with respect to the distribution of the measured data $\mathcal{U}$. The first term in the decomposition of the MSE is the bias term. The less restrictions imposed on $\hat{\theta}$, the best the estimate can represent the true parameter θ and hence the smaller its bias. The second term is the variance term, which is smaller the more constrained the estimate $\hat{\theta}$ is (less freedom to adjust to the data).

While we only outline the bias-variance decomposition for the MSE criterion, the distortion-variability trade-off is more general. Restrictions introduced in model fitting procedures reduce their sensitivity to noise but potentially introducing distortions to the fitted model. The exact balance between distortion (bias) and variability (variance) depends on many factors such as the noise level and the number of observations available, among others. When modal estimation is made as described above, the appropriate order q of the AR model for the acceleration process is chosen based on stability plots and can be interpreted as a balancing measure between bias and variance as above.

Formulation of the banded AR(q) model

The implicit set of restrictions in a AR(q) model consists of setting all $\beta_{i_1,i_2,j}$ coefficients to zero for $j > q$. Such restrictions can be interpreted as an assumption that observations in the too far past should not have an effect on present observations. The restrictions we propose can be interpreted as the spatial counterpart of the restrictions made by the AR process in time: they correspond to an assumption that the signals in nodes that are far apart do not have a direct effect on each other, or that their effect is accounted for by the closer nodes.

One complication with introducing such restrictions is that, differently from time, different metrics can be used to represent the distance between two elements in space. To define a suitable metric measuring distance between acceleration measurements in the structure, the distance $d(i_1, i_2)$ between nodes i_1 and i_2 is defined as the minimum number of hops needed to travel between the two nodes. Just as the AR(q) sets $\beta_{i_1,i_2,j} = 0$ if $j > q$, we use the restriction $\beta_{i_1,i_2,j} = 0$ if $d(i_1, i_2) > w$ for a certain threshold $w > 0$. Figure 1 shows a structure with linear topology, along with the corresponding connection graph and the hop-distance between nodes. The restricted AR(q) model is then represented by:

$$\ddot{\mathbf{u}}(k) = \sum_{j=1}^{\infty} \mathbf{L}_j^{(q,w)} \ddot{\mathbf{u}}(k-j) + \upsilon(k), \tag{3}$$

where for each matrix of coefficients $\mathbf{L}_j^{(q,w)}$, $\beta_{i_1,i_2,j}^{(q,w)} = 0$ whenever $d(i_1, i_2) > w$ or $j > q$. For a linear topology structure as shown in Figure 1, these restrictions result in banded coefficient matrices. Setting $w = 1$, for instance, leads to the matrices of coefficients of the form:

$$\mathbf{L}_j^{(q,1)} = \begin{cases} \begin{bmatrix} \beta_{1,1,j}^{(q,w)} & \beta_{1,2,j}^{(q,w)} & 0 & 0 & 0 \\ \beta_{2,1,j}^{(q,w)} & \beta_{2,2,j}^{(q,w)} & \beta_{2,3,j} & 0 & 0 \\ 0 & \beta_{3,2,j}^{(q,w)} & \beta_{3,3,j}^{(q,w)} & \beta_{3,4,j}^{(q,w)} & 0 \\ 0 & 0 & \beta_{4,3,j}^{(q,w)} & \beta_{4,4,j}^{(q,w)} & \beta_{4,5,j}^{(q,w)} \\ 0 & 0 & 0 & \beta_{5,4,j}^{(q,w)} & \beta_{5,5,j}^{(q,w)} \end{bmatrix}, & \text{for } j = 1, \ldots, q \text{ and,} \\ \mathbf{0}, & \text{if } j > q. \end{cases} \tag{4}$$

As the $L^{(q,w)}$ notation suggests, the problem of selecting a proper bandwidth w is akin to that of selecting a proper number of lags q for the autoregressive model.

Communication requirements

The estimation of the coefficients in the full AR process requires that each node compute its correlations with all other nodes up to lag q. For the restricted model, on the other hand, computing the column of coefficients associated with node i_2 only requires the correlations with the nodes i_1 such that $d(i_1, i_2) \leq w$ up to correlation q. Suppose that each node transmits its observed acceleration to all other nodes who need it to compute the estimates for its AR parameters. The transmission volume over the wireless network can be computed based on the number of sample-hops that is necessary to estimate the model.

For a network with a linear topology the signal at each node must be transmitted to all other nodes in the network. The total transmission volume in this case is $np(p+1)/2$ with order $O(np^2)$, where p is the number of the nodes in the network and n is the number of samples at each node.

For the restricted model, each node has to communicate its data over a graph with diameter $2 \cdot w$, with a transmission volume of

$$n\left\{2w\left(p-2w\right)+2\left[\left(2w-1\right)+\left(2w-2\right)+...+\left(2w-w\right)\right]\right\}=2wp-w^2-w$$

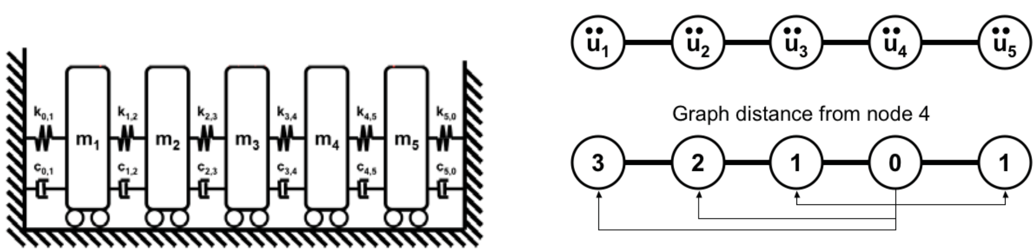
or $O(nwp)$. This corresponds to an efficiency ratio of $O(w/p)$, which could be a significant saving in communication load of the network.

Simulated Examples

In this section a set of simulated examples are used to evaluate the performance of the banded AR model by comparing the rate of convergence of the modal properties of the system in the banded versus full models. The simulation parameters are described first and then the results are presented.

Simulation Set-up

Two structures with linear topology and lumped masses are used to simulate data. Figure 1(a) shows the schematic of one of the structures with five lumped masses and linear stiffness and damping components connecting them. The second simulated case has a similar structure, but with ten lumped masses instead of five. In each case the acceleration response of the structure to white noise excitation is simulated at each mass and the performance of the banded and full AR(q) models are compared for different neighborhood width w for the restricted coefficient matrices. Stabilization graphs, with the lags ranging from 5 to 40 are used to show the rate of convergence of the results [6]. Cases 1 and 2 correspond, respectively, to the five and ten degree of freedom structures whose modal parameters are listed in Table I and mode shapes as shown in Figure 2.



(a) Linear topology structure with 5 nodes (b) Hop distance from Node 4

Figure 1: Structure with Linear Topology. Each node in the structure is only connected to its closest neighbors.

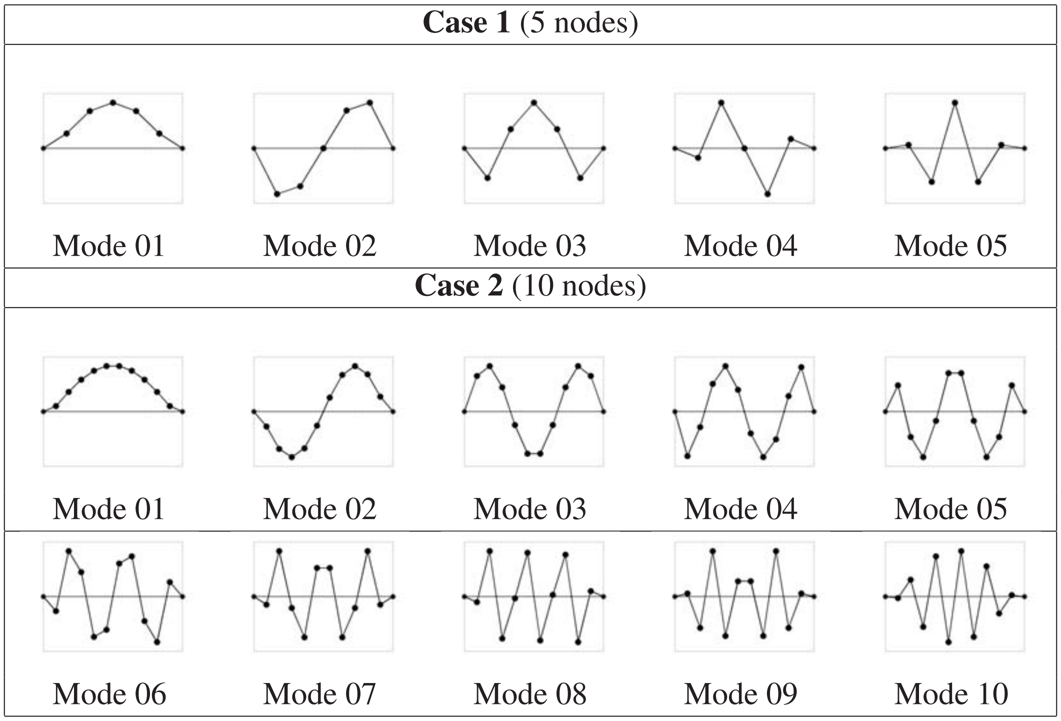


Figure 2: Mode shapes for the simulated structures in Cases 1 and 2.

Case 1: Five Degrees of Freedom

The stabilization diagrams obtained with number of lags varying from 5 to 40 and for neighborhood widths w varying from 0 to 2 as well as for the full model are shown in Figure 3. The graphs show that even when disregarding the cross-correlation coefficients ($w = 0$), the frequencies are identifiable. More importantly, the stabilization graph for the case where $w = 1$ shows a rapid convergence of all of the five estimated modes. The communication load in this case is only 53% of the full model, which provides a great saving. The stabilization graphs for the case where $w = 2$ and the full model are practically identical. In this case, however, the communication load is 93% of the full model, which is not a significant saving.

Table I: MODAL PARAMETERS FOR THE SIMULATED STRUCTURES.

	Mode Number	1	2	3	4	5	6	7	8	9	10
Case 1	Frequency (Hz)	3.0	4.4	5.3	7.3	9.1	-	-	-	-	-
	$\zeta_j(\%)$	5.0	4.0	3.7	3.6	3.7	-	-	-	-	-
Case 2	Frequency (Hz)	1.6	3.0	4.0	4.7	5.4	6.5	7.6	8.5	9.2	9.6
	$\zeta_j(\%)$	5.0	4.0	4.1	4.3	4.6	5.1	5.7	6.2	6.5	6.8

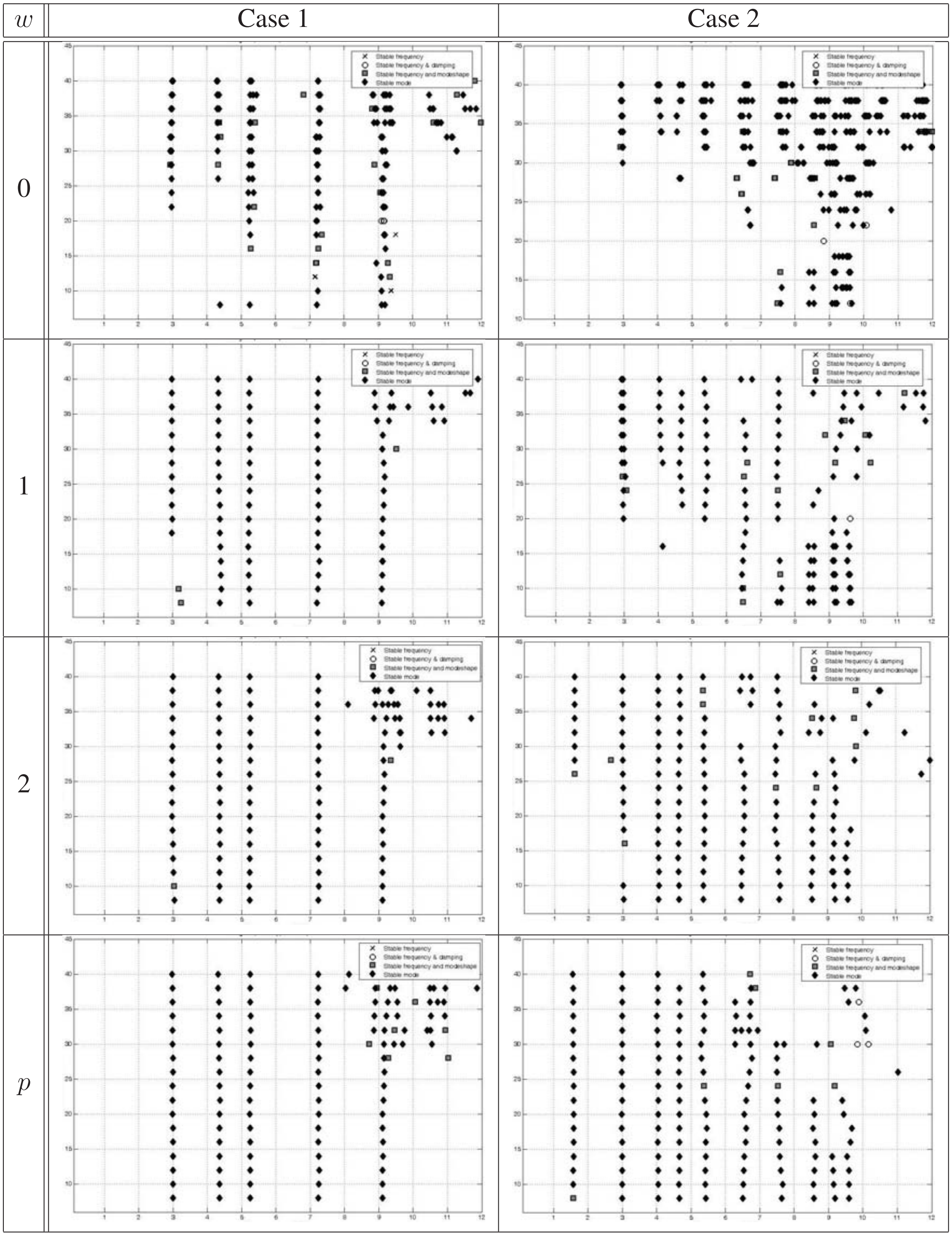


Figure 3: Stabilization plots using different neighborhood sizes: model orders (number of lags) are in the vertical axis and frequencies (in Hz) are in the horizontal axis. Full black diamonds indicate stable modes, gray squares indicate stable frequencies and mode shapes, circles indicate stable frequencies and damping ratios and crosses indicate stable frequencies only.

Case 2: Ten Degrees of Freedom

Similar stabilization diagrams for the ten-DOF case are also shown in Figure 3. In this case the first mode is not identified with an AR model with 40 lags in the case where $w = 1$, but the rest of the modes are identified. Note that the communication load in this case is only 33% of the full model. The stabilization graphs for the case where $w = 2$ show that all of the ten modes are identified. In this case the communication load is 62% of the full model, which again is a significant saving.

Conclusion

Multivariate auto-regressive (AR) models are commonly used to model the dynamic behavior of vibrating structures and to infer the modal parameters of a structure subject to ambient vibration. In this paper, a restriction regime to AR models is presented that can significantly reduce the volume of transmitted data over a wireless sensor network while introducing little distortion on the modal parameter estimates.

Simulated data of structures with linear topology are used to show that the restricted models can recover the modal parameters while reducing the volume of transmitted data.

Future work will be devoted to proposing a methodology for selecting an appropriate set of restrictions to the AR models and the analysis of the savings that can be achieved in structures with alternative topologies.

References

[1] Doebling, S. W., C. R. Farrar, and M. B. Prime. MArch 1998. “A Summary Review of Vibration-Based Damage Identification Methods”, *The Shock and Vibration Digest*, 30(2):91–105.

[2] Pandit, S. M. 1991. *Modal and Spectrum Analysis: Data Dependent Systems in State Space*, John Wiley and Sons.

[3] Roeck, G. D., W. Claesen, and P. V. D. Broeck. 1995. “DDS-Methodology Applied to Parameter Identification in Civil Engineering Structures”, *in Proceedings of Vibration and Noise*, Venice, Italy, pp. 341–353.

[4] Pakzad, S. 2008. *Statistical Approach to Structural Monitoring Using Scalable Wireless Sensor Networks*, Ph.D. thesis, UC Berkeley Civil and Environmental Engineering Department.

[5] Bickel, P. J. and B. Li. 2006. “Regularization in Statistics”, *Test*, 15(2):271–344.

[6] Pakzad, S. and G. Fenves. 2009. “Design and Implementation of Scalable Wireless Sensor Network for Structural Monitoring”, *ASCE Journal of Infrastructure Engineering*.