A Bayesian Network Framework for Post-earthquake Infrastructure System Performance Assessment

M.T. Bensi,[1] A. Der Kiureghian, [1] and D. Straub[2]

[1]University of California at Berkeley, Department of Civil and Environmental Engineering. Berkeley, CA 94720 (mbensi@berkeley.edu, adk@ce.berkeley.edu)
[2]Technical University of Munich, Engineering Risk Analysis Group, 80290 München, Germany (straub@era.bv.tum.de)

ABSTRACT

We are currently working toward the development of a probabilistic decision-support system (DSS) for near-real time emergency response and recovery following a seismic event. We utilize a Bayesian Network (BN) framework for assessing the performance of spatially distributed infrastructure systems subjected to seismic hazards. BNs have many valuable characteristics (e.g. they are graphical, adaptable, efficient, and enable Bayesian updating) that make them well-suited for such an application. We model seismic demands on an infrastructure system by constructing a BN model of ground motion intensity as a spatially distributed random field. In this paper we present this seismic demand model with recent enhancements including the addition of a finite source model. The performance of infrastructure system components and the overall system is modeled using fragility functions and BN system connectivity formulations. An example application to a hypothetical transportation system is presented to illustrate the power and value of the proposed BN framework.

INTRODUCTION

Civil infrastructure systems are the lifelines of modern societies. Disruption of the service of critical infrastructures such as transportation networks, utility distribution systems, communication networks, as well as government, health, and public facilities can have far-reaching economic, social, health, and safety consequences. Meanwhile, infrastructure systems remain vulnerable to a wide variety of natural and manmade hazards, and it is required to study and optimize their performance under these hazards. In this paper we focus on the earthquake hazard, which is a dominant hazard to infrastructure systems in many areas of the world.

For many hazards, including earthquakes, the occurrence of the event cannot be prevented, but it is possible to mitigate hazards by minimizing the consequences when an event does occur. Decisions made by emergency management personnel immediately after a major earthquake (e.g. deployment of emergency personnel and equipment, evacuation of people, post-event inspections, and closure or continued operation of critical facilities) can significantly influence the severity of the associated consequences.

Unfortunately, these vital decisions are often made in an ad hoc manner, under large uncertainties, with information that evolves in time. As a result, post-event decision-makers are in need of a tool that (1) can be employed in near-real time following a major disaster, (2) is capable of synthesizing evolving incoming information, and (3) can properly account for uncertainties both in the incoming information as well as in the analytical models that are used to assess the states of various components of an infrastructure. Recent advances in technology, computer science, risk assessment, and hazard modeling have yielded knowledge and methodologies that can be integrated to create a unique tool to aid the post-event decision-making process.

To address this need, we are currently working towards the development of a probabilistic decision-support system (DSS) for near-real time emergency response and recovery following a seismic event, utilizing a Bayesian Network methodology. This DSS will integrate information from a wide-range of sources in near-real time to provide a comprehensive description of the state of a geographically distributed infrastructure system following an earthquake in near real time.

A Bayesian Network (BN) (also known as a Belief Network or Probabilistic Network) is a probabilistic graphical model that encodes a set of random variables and their probabilistic dependencies (Jensen & Nielson 2007). The BN methodology has been employed because it has several characteristics that are ideally suited for the proposed application: (1) The BN is a graphical, powerful, and efficient tool for representing systems having components with uncertain demands and capacities, (2) BNs can be used to model multiple hazards and their interdependencies, (3) BNs can be used to model interacting systems, (4) BNs provide efficient frameworks for probabilistic updating and the assessment of component/system performance, (5) BNs can be used to identify critical components and cut sets in a system, (6) the graphical interface makes it an excellent tool for use by practitioners and end-users, and (7) BNs can be extended to include utility and decision nodes and thus analysis can be performed in which decision alternatives are considered and ranked based on expected utilities.

The evaluation and management of seismic risks posed on an infrastructure system requires a number of elements, including: (1) characterization of the earthquake magnitude, location and other source features, (2) estimation of ground motion intensities at distributed points in the system while properly accounting for the spatial correlation structure, (3) modeling of the performance of system components, e.g., in terms of fragility functions, (4) modeling of the performance of the system in terms of the component performances and the seismic demand, and (5) system reliability assessment under different earthquake scenarios. In the context of the aforementioned near-real time DSS, we are interested in the reliability of the system and its components conditioned on any information that may become available following an earthquake, e.g. known magnitude and location of the earthquake, measurements of shaking intensity at selected locations, observed performance of selected components. This paper describes the implementation of these requirements in the BN framework.

SEISMIC DEMAND MODEL

Probabilistic seismic hazard analysis (PSHA) has grown as a computational tool in civil engineering in recent years, though it is primarily used for design and analysis of single-site structures or systems. The primary objective of PSHA is to determine the probability of exceeding a specified level of earthquake-induced ground shaking at a specific site. For a given earthquake event (of random magnitude and location), this probability can be written as

$$P(S_i > s_i) = \sum_{\text{all sources}} \iiint_{r_i, m, e_i} P(S_i > s_i | m, r_i, e_i) f_{M,R_i}(m, r_i) f_{\epsilon_i}(e_i) dm dr_i de_i \qquad (1)$$

where S_i denotes the intensity of ground shaking at site i, M and R_i represent the random earthquake magnitude and source-to-site distance, respectively, and ϵ_i represents a standard error term associated with the regression model that predicts the intensity at the site for given earthquake magnitude and distance.

Existing research in PSHA primarily focuses on systems located at individual sites. Civil infrastructure systems, on the other hand, tend to be geographically distributed. As such, they possess certain distinguishing features relative to single-site systems, which should be carefully considered. These include:

1. Geographically distributed systems may be exposed to a variety of seismic hazards such as ground shaking, fault rupture, land/rock-slide, and liquefaction, whereas an individual site may only be exposed to a select few of these hazards. Furthermore, because it covers a larger area, a spatially distributed system has a higher rate of exposure to earthquake events than a single-site system. While we focus solely on ground-shaking hazard in this paper, it is important to note that the BN framework allows extension to include other hazards and this remains an area for future study.
2. Various attributes of seismic hazard for a spatially distributed system may be correlated in space. For example, intensities of ground shaking affecting components of an infrastructure system at two different locations may be correlated. To properly account for this correlation, such attributes of an earthquake must be modeled as random fields.

Within the context of modeling seismic demands on a spatially distributed infrastructure system, one is interested in the probability that ground motion intensities at any subset j of one or more sites i exceed corresponding thresholds:

$$\Pr\left(\bigcup_j \left(\bigcap_i \{S_i > s_i\}\right)\right) \qquad (2)$$

With this objective in mind, we next discuss the BN model for predicting the seismic intensities at different sites across a spatially distributed infrastructure system following an earthquake. The intensity of ground motion at a site is often characterized in terms of a single measure, such as the peak ground acceleration or the spectral acceleration at a selected frequency. As defined previously, let S_i denote this intensity measure at site i. Predictive models based on regressions of observed data are available to relate S_i to site and earthquake source characteristics (Bozorgnia and Bertero, 2004). These models typically have the form

$$\ln(S_i) = f(M, R_i, \mathbf{X}_i) + \epsilon_{r,i} + \epsilon_m \qquad (3)$$

where $f(M, R_i, \mathbf{X}_i)$ is a deterministic function of the magnitude (M), site-to-source distance (R_i), and other characteristics of the source and site ($\mathbf{X}_i$), such as the type of faulting mechanism and the site shear-wave velocity; ϵ_m is an inter-event, zero-mean normally distributed error term representing inaccuracies in source characterization, and $\epsilon_{r,i}$ is a zero-mean normally distributed intra-event error term representing inaccuracies due to site-specific factors and wave propagation effects. In general, the error terms $\epsilon_{r,i}$ and $\epsilon_{r,j}$ at two sites i and j are correlated. Among other authors, Boore et al. (2003) and Park et al. (2007) have investigated this effect and have developed autocorrelation models expressed as functions of the inter-site distance. The BN representing the above model of ground motion intensities and assuming a point-source earthquake model is shown in Figure 1. Note that due to the spatial correlation described above, all nodes representing the intra-event error terms are connected through links. Straub et al. (2008) discuss the computational complexities (and potential computational intractability) associated with such a formulation. To overcome this difficulty, Straub et al. (2008) employed a principal-component formulation whereby the correlated variables $\epsilon_{r,i}$ are approximately described in terms of a few standard normal variables $u_i, i = 1,2, \ldots, m$ with $m \ll n$. The corresponding BN model is shown in Figure 2.

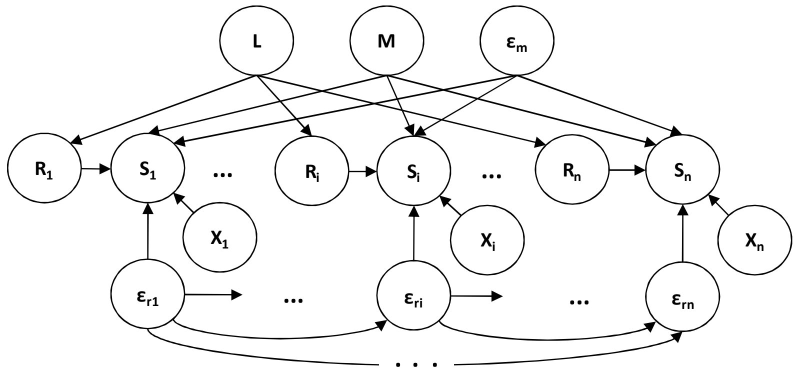


Figure 1: BN of seismic demands with spatially correlated inter-event error terms

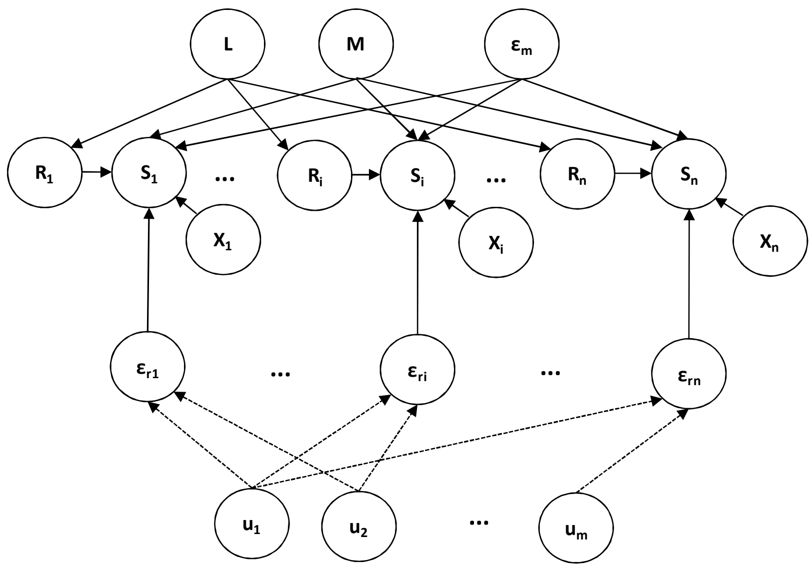


Figure 2: BN of seismic demands using approximation of spatial correlation

It is seen in Figure 2 that the ground motion intensity at site i, S_i, is a function of the magnitude, source-to-site distance, and error terms as indicated by the directed links going into S_i from its parent nodes M, R_i, $\epsilon_{r,i}$, and ϵ_m. $\epsilon_{r,i}$ is a function of the statistically independent standard normal random variables (u_i) as detailed in Straub et al. (2008). The source-to-site distance is a function of the earthquake location (L) since a point source model has been assumed in this figure.

Earthquakes, however, do not occur as points on a fault, but rather as finite ruptures along a fault length, area, or volume. Thus we expand the model to include a finite source model. Figure 3 shows a conceptual framework for including a finite source model in the BN. The elements of this model are as follows. First, we express the rupture length as a function of the earthquake magnitude and a standard error term using available empirical models such as in Wells and Coppersmith (1994):

$$\log(L_{rup}) = a + bM + \epsilon \tag{4}$$

Here, L_{rup} is the rupture length, M is the earthquake magnitude, a and b are regression constants, and ϵ is a standard error term.

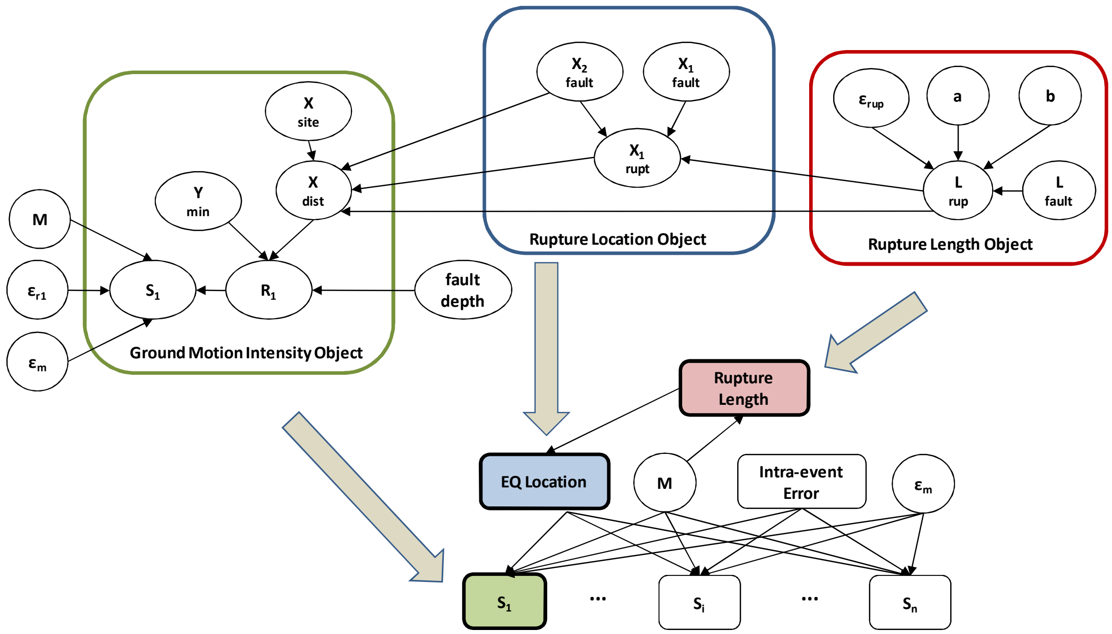


Figure 3: Conceptual framework of BN including finite source model

In Figure 3 we are making use of objects to reduce graphical clutter. In the BN framework an object is a special node behind which another BN resides. Objects do not change the topology (variable dependencies) of the BN; an object simply hides some nodes to make it easier to read the BN. Thus, we introduce an object (on the top right in Figure 3) to model the rupture length as a function of magnitude.

Next, we use the geometry of the seismogenic source to describe the location of the rupture along the fault. The left edge of the rupture ($X_{1,rupt}$) is a random variable distributed uniformly (or by any other distribution of choice) on the interval $[X_{1,fault}, X_{2,fault} - L_{rup}]$, where $X_{1,fault}$ and $X_{2,fault}$ are the left and right edges of the fault relative to a reference coordinate system. These relations are encoded in the BN through the earthquake location object (top middle).

We must update the ground motion intensity object in Figure 3 to calculate the source-to-site distance for the extended source model. Based on the coordinate system and notation defined in Figure 4, with the fault idealized as a straight line, the x-component of the shortest distance vector from the site to the rupture is calculated as $\max[X_{1,rupt} - X_{site}, X_{site} - \min[X_{1,rupt} + RL, X_{2,fault}], 0]$. This value together with the distance in the direction perpendicular to the fault (Y_{min}) and the depth to the rupture plane define the shortest distance from the site to the rupture. These relations are encoded in the ground motion intensity object of the BN through four additional nodes as shown in Figure 3 (top left).

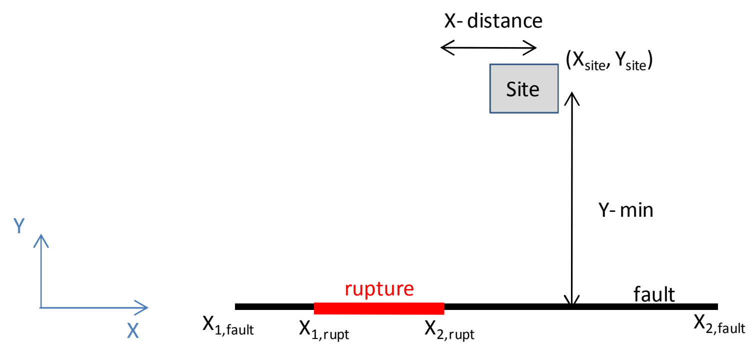


Figure 4: Parameters for defining source-to-site distance

SYSTEM PERFORMANCE

Conditioned on the distribution of ground motion intensity at its site, the performance of an individual infrastructure system component is modeled using fragility functions. Fragility functions provide the probability of damage given a level of ground motion intensity. Based on the performance of its components, system performance can then be modeled using system connectivity formulations provided in Bensi et al. (2009). The resulting conceptual BN in which component and system performances are included is presented in Figure 5. Nodes c_i denote component states and they are dependent on the corresponding ground motion intensities as well as site-specific factors. The system node is dependent on the component states. This BN formulation is referred to as a "conceptual BN" because modeling system connectivity directly as shown in Figure 5 may not be computationally efficient or even feasible. To make the BN computationally more tractable, a more sophisticated system formulation must be used, as described in Bensi et al. (2009).

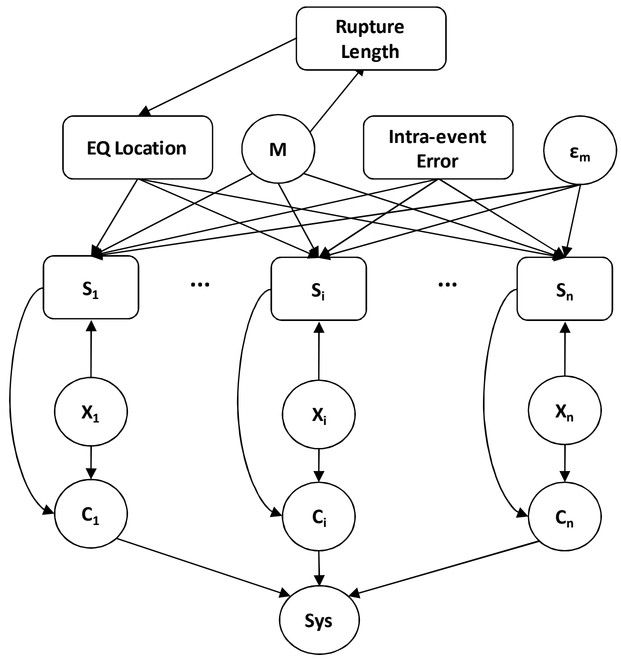
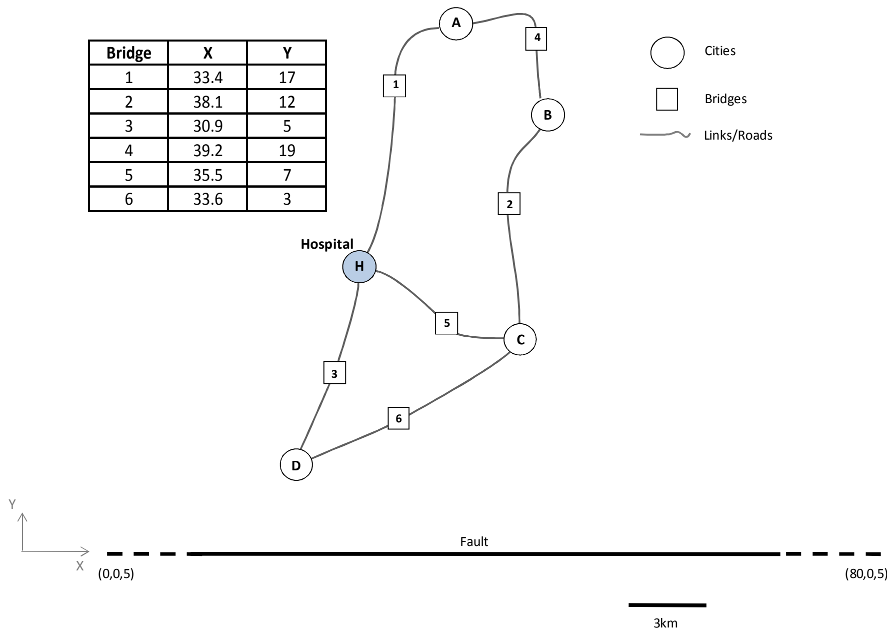


Figure 5: Conceptual BN including seismic demand model and component/system performance

With the seismic demand and system performance models, we are able to assess the state of our infrastructure system following an earthquake while incorporating available information. That is, we can enter evidence or observations at any of the nodes, and this information will propagate through the network to update our beliefs about all other variables in the BN. For example, we can determine component and system failures probabilities given that we know an earthquake has occurred as well as its magnitude and location. Additionally, via the max-propagation algorithm (Jensen & Nielson 2007; Friis-Hansen 2004), we can determine likely component configurations given a specified set of evidence. Such information will be useful for post-event applications such as determining components that are most likely to have failed given a set of evidence (e.g. to aid the dispatch of inspectors) or pre-event for determining weak links in the system. The following example demonstrates these notions more concretely.

EXAMPLE

Consider the hypothetical transportation network shown in Figure 6. In this system, a set of cities (circles) are connected to a hospital via roadways that cross bridges (squares) that can fail during an earthquake. We are interested in the probability that the hospital will be accessible from each city following an earthquake (i.e. system failure is defined as the event that any of the cities is unable to reach the hospital). The BN for modeling this system is shown in Figure 7. The empirical and analytical models that have been employed in this example are summarized in Table 1. The choice of models used in this example is based on their characteristics or ease of applicability and presentation. They should in general be viewed as “place-holders” rather than restrictive choices. It is fairly easy to switch one model for another within the BN.



Bridge	X	Y
1	33.4	17
2	38.1	12
3	30.9	5
4	39.2	19
5	35.5	7
6	33.6	3

Figure 6: Hypothetical transportation system

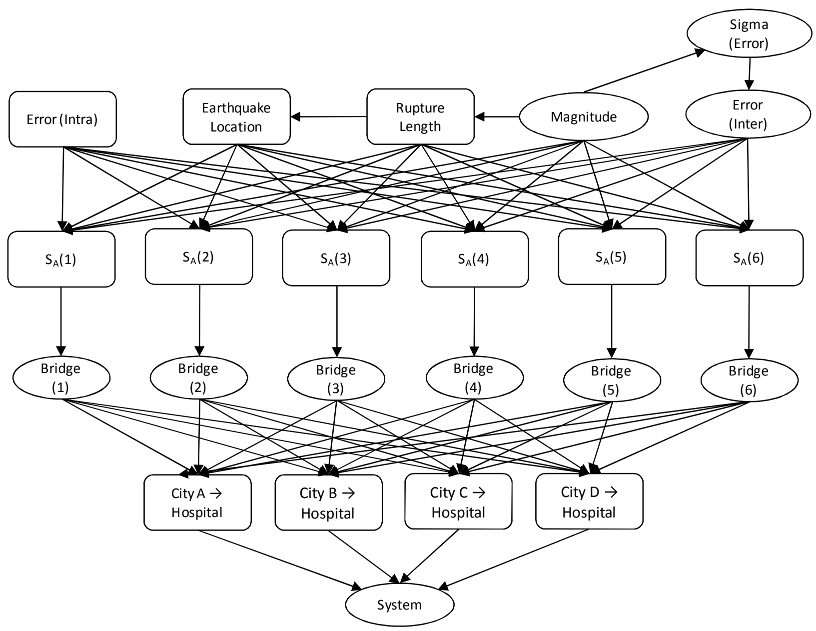


Figure 7: Example System BN

Table 1: Empirical/analytical models employed in example BN

magnitude distribution	Truncated GR (exponential)
ground motion (spectral acceleration at T=1.0 sec) prediction equation	Campbell (1997)
magnitude-rupture length relation	Wells & Coppersmith (1994)
spatial error correlation	Boore et al. (2003), Park et al. (2007)
error approximation model	Straub et al. (2008)
fragility function	adaptation of Gardoni et al. (2002)
system connectivity	Bensi et al. (2009)

To demonstrate the notion of information updating, we consider several post-earthquake scenarios:

(1) An earthquake has occurred, but we have no information available about its magnitude, location or other characteristics, i.e. we have no evidence.
(2) We observe that the earthquake magnitude is in the range 6.25-6.5 and that a 15-20km rupture occurs on the far left edge of the fault.
(3) We observe that the earthquake magnitude is in the range 6.25-6.5, a 15-20km rupture occurs on the far left edge of the fault, and bridge 1 has failed.

Table 2 demonstrates how available information can update our assessments of component and system performance. For example, consider the case in which we have only observed the magnitude, rupture length, and location of the earthquake versus the case in which we observed the same earthquake characteristics information but also that bridge 1 has failed. The information that the bridge has failed back-propagates through the network by updating our belief about the distribution of spectral acceleration at the site of bridge 1, then updating the inter- and intra-error structure and back down to update the distribution of spectral accelerations at the other sites in the network and eventually their respective bridge failure probabilities. That is, the change in the posterior probability of failure of another bridge (e.g. bridge 2), from the case without observation of bridge 1 failure to the case that includes the observation, is due to the correlation in the ground motion intensities arising from the inter- and intra-error terms.

Additionally, we can use the max propagation algorithm to determine the most likely configuration of component states for any given evidence. The numbers listed in Table 3 indicate the relative likelihoods of component states (failure or survival) given the aforementioned evidence scenario (2) and the information that the system has failed, i.e. one or more cities have been unable to reach the hospital. That is, values of 1.0 in this table indicate that a particular component state is part of the most probable configuration of all components states in a network given the evidence. Thus, given evidence scenario (2) and failure of the system, the most probable configuration of the system is bridges 3 and 6 have failed and all other bridges have survived. This information indicates bridges 3 and 6 are most critical for the survival of the system and that these bridges should be prioritized when planning mitigation action.

Table 2: Bridge and system failure probabilities for example system

Evidence Case	Bridge 1	Bridge 2	Bridge 3	Bridge 4	Bridge 5	Bridge 6	System
Unconditional (no evidence)	0.003	0.004	0.006	0.003	0.006	0.006	0.001
Magnitude = 6.25-6.5, Rupture length = 15-20km, Location=Left edge of fault	0.013	0.013	0.039	0.004	0.022	0.031	0.004
Magnitude = 6.25-6.5, Rupture length = 15-20km, Location=Left edge of fault, bridge 1 has failed	1.000	0.039	0.066	0.010	0.054	0.053	0.065

Table 3: Likely configuration of bridges given evidence case (2) and known system failure

	Bridge 1	Bridge 2	Bridge 3	Bridge 4	Bridge 5	Bridge 6	System
failure	0.85	0.85	1.00	0.04	0.49	1.00	1.00
survival	1.00	1.00	0.85	1.00	1.00	0.85	0.00

The results demonstrated in this example are just a small subset of the possible evidence cases that can be considered using the BN that has been developed. It is hoped that they demonstrate the power of the BN framework for near-real time applications.

CONCLUSIONS

The goal of this paper was to present some of our recent efforts devoted to the development of BN framework for seismic infrastructure system performance assessment and near-real time decision support. A seismic demand model was presented in which ground motion intensities at different sites across a distributed infrastructure system are predicted by modeling the spatially distributed random field using BN. Conditioned on the distribution of ground motion intensity at each site, component performance is modeled using available fragility functions. System performance is modeled using formulations available in Bensi et al. (2009). The BN framework allows for information updating so that system performance can be assessed in light of all available information. An example of a hypothetical transportation system was presented to make ideas outlined in this report more concrete.

ACKNOWLEDGEMENT

The first author acknowledges support from a National Science Foundation Graduate Research Fellowship. Funding for this research was received from the PEER Lifeline program, which is gratefully acknowledged.

REFERENCES

Abrahamson, N.A., 2000. Effects of rupture directivity on probabilistic seismic hazard analysis. In *Proc. of the 6th International Conference on Seismic Zonation, Palm Springs, Earthquake Engineering Research Institute*.

Bensi, M. et al., 2009. Modeling infrastructure system performance using BN. In *Proc. of the 10th International Conference on Structural Safety and Reliability*. Osaka, Japan: Accepted.

Boore, D.M. et al., 2003. Estimated Ground Motion From the 1994 Northridge, California, Earthquake at the Site of the Interstate 10 and La Cienega Boulevard Bridge Collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America*, 93(6), 2737-2751.

Bozorgnia, Y. & Campbell, K., 2004. *Chapter 5: Engineering Characterization of Ground Motion.* In: *Earthquake Engineering From Engineering Seismology to Performance-based Engineering*, Boca Raton, Florida: CRC Press.

Campbell, K., 1997. Near-Source Attenuation Relationships for Horizontal and Vertical Components of Peak Ground Acceleration, Peak Ground Velocity, and Pseudo-Absolute Acceleration Response Spectra. *Seismological Research Letters*, 68(1), 154-179.

Friis-Hansen, P., 2004. Structuring of complex systems using Bayesian Networks (with additional material from the unpublished extension to the paper, 2008). In O. Ditlevsen & P. Friis-Hansen, eds. *Two Part Workshop at DTU*. Technical University of Denmark.

Gardoni, P., Der Kiureghian, A. & Mosalam, K.M., 2002. *Probabilistic Models and Fragility Estimates for Bridge Components and Systems - PEER Report 2002/13*, University of California, Berkeley.

Jensen, F.V. & Nielson, T.D., 2007. *Bayesian Networks and Decision Graphs* second., New York: Springer Science Business Media|Springer.

Park, J., Bazzurro, P. & Baker, J.W., 2007. Modeling spatial correlation of ground motion Intensity Measures for regional seismic hazard and portfolio loss estimation. In T. Kanda, ed. *Applications of Statistics and Probability in Civil Engineering*. Tokyo, Japan: Taylor & Francis Group.

Somerville, P.G. et al., 1997. Modification of Empirical Strong Ground Motion Attenuation Relations to Include the Amplitude and Duration Effects of Rupture Directivity. *Seismological Research Letters*, 68, 199-222.

Straub, D., Bensi, M. & Der Kiureghian, A., 2008. Spatial Modeling of Earthquake Hazard and Infrastructure System Performance Through Bayesian Network. In *Proc. Inaugural International Conference of the Engineering Mechanics Institute (EM08)*. University of Minnesota, Minneapolis.

Wells, D. & Coppersmith, K., 1994. New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement. *Bulletin of the Seismological Society of America*, 84(4), 974-1002.