Modelling of the node immunity change process in a network system

I. Žutautaitė-Šeputienė[1,2], J. Augutis[1,2] & E. Ušpuras[1]
[1]*Lithuanian Energy Institute, Lithuania*
[2]*Vytautas Magnus University, Lithuania*

Abstract

The processes of network nodes' resistance/immunity to hazard spread were analyzed in this study. The case when nodes' immunity is non-constant (which depends on various factors) is analyzed. The mathematical algorithm (based on the Bayesian approach) for non-constant network nodes' immunity updating by observed data is presented in this paper. A numerical experiment was performed to illustrate the application of the developed algorithm.

Keywords: Bayesian approach, non-constant immunity modelling.

1 Introduction

More and more attention is focused on the problems related to quantitative assessment of the behaviour of systems/components/infrastructures in terms of dependability and security and, more generally, quality of service indicators. People, technical equipment, computers (hardware, software), etc or their various combinations form a structure network system (i.e. nodes and network lines/channels) (fig. 1).

A hazard/error, its propagation and negative effect, the natural physical ageing of the system, the influence of external factors to the system vitality and so on affect the network system dependability and quality of the service. On the other hand, in a network system some (or all) nodes can have immunity from hazards (for example, in computers we have firewalls, antivirus programs against hacker attacks and computer viruses), according to the hazard propagation immunity of the system node changes (which could increase).

When analyzing network systems under the influence of the hazard propagation process, one point of interest is the modelling of node resistance

WIT Transactions on Information and Communication, Vol 39, © 2008 WIT Press
www.witpress.com, ISSN 1743-3517 (on-line)
doi:10.2495/RISK080351

(immunity changes). A numerical experiment of the proposed algorithm application for updating the node immunity function is presented in this study.

This study could be used for:

- Forecasting how many cycles are necessary to eliminate the hazard or to reduce it to a safe level; how long (how many cycles) the system will be able to work normally under the influence of the hazard.
- Modelling of the hazard distribution in the network with non-constant node immunities.

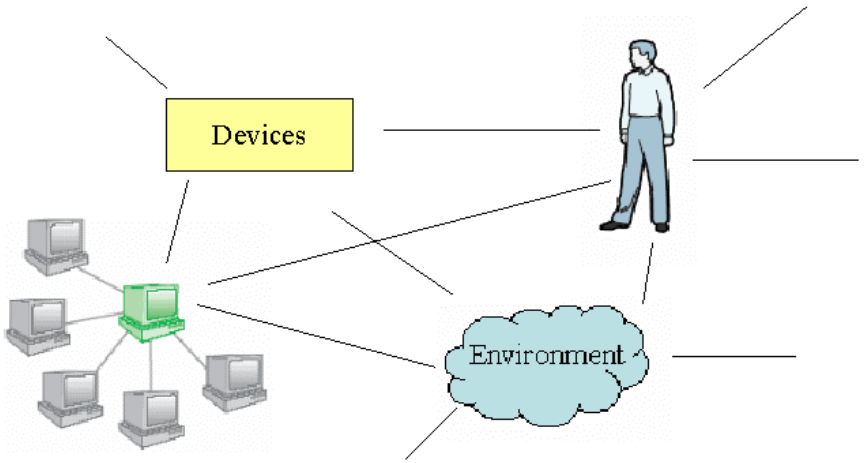

Figure 1: An example of a general network system.

2 Definitions of hazard, immunity, flows and others concepts

Hazard is defined as a feature or characteristic of a material, a technological process, information, the negative effect of human activities, the natural physical ageing of the system, the influence of external factors or other phenomena that specifies a potential possibility of endangering human life, health, nature, buildings, equipment, etc. Illustrations of hazard could be poisonous chemical substances kept in stock, an open source of radiation, a car with brakes out of order, a lock that is built on a river that flows through a city, etc. Hazard measurement is less clearly defined; therefore here such qualitative evaluations as high hazard, medium hazard or low hazard are used. Nevertheless, in certain cases, quantitative expressions are also used. For example, when we speak about a chemical material with definite toxicity, the quantity of that material is an important characteristic. On the internet, one of the computer risk characteristics is a number of computer viruses, etc. Sometimes, certain equivalents are used, for example, explosive materials are compared to the trotyl equivalent.

The development and expansion of various networks and network structures create favourable conditions for the spread of hazard through network channels.

Examples of such phenomena could be the spread of toxic/radioactive substances or high temperatures; and its influence on devices, people, etc.

Immunity/resistance is the systems' ability to resist negative effects (hazard). Illustrations of immunity could be antivirus programs, firewalls against a computer virus, hacker attacks, fire protection in buildings, traffic lights at crossroads, the immune system of an organism, etc.

Now we will define several concepts that will be used in the paper.

The number of network nodes. The number of network nodes is marked as N.

Additive hazard. This is a type of hazard where hazards in the nodes of the network can be added to or a part of the hazard moved to the other nodes. Examples of the additive hazard are: the collection of hazardous materials, the amount of fake money in the supermarkets, transport intensity at the crossroads, etc, marked as H.

Flow intensity. q_{ij} is the coefficient of flow intensity in the network lines or channel; it marks the part of the hazard in the node i that will be transmitted to the node j. The intensity of the flow to the node j and the intensity of the flow from the node j are defined respectively

$$\breve{q}_j = \sum_{\substack{i=1 \\ i \neq j}}^{N} q_{ij}\,,\ \hat{q}_j = \sum_{\substack{k=1 \\ k \neq j}}^{N} q_{jk} \tag{1}$$

Hazard transfer cycle. Hazard transfer in the network from one node to the other is regarded as one hazard transfer cycle.

Network node immunity. $I_j(\cdot)$ is the coefficient of the network node immunity. It marks which part of the hazard is stopped, before getting into the node j ($0 \leq I_j \leq 1$, i.e. percentage). Node immunity can be created by the security systems, anti-virus computer software, etc. The "observed" value of node j immunity could be obtained

$$I_{ji} = \frac{\breve{q}_{ji} - P_{ji}}{\breve{q}_{ji}}, \tag{2}$$

$i = 1, 2, \ldots$, here i is the number of cycles, $\breve{q}_{ji}$ is the flow of hazard to the node j in the i^{th} cycle, P_{ji} is the amount of hazard that gets into the node j during the i^{th} cycle.

An example of a simple three node network system with describing elements is presented in fig. 2.

The mechanism of hazard propagation in network systems in the case of a single hazard evolved in one of the network nodes and in the case when hazard arises during each cycle (with constant node immunities) was analyzed by Augutis et al [1]. In their study the marginal hazard distribution mathematical model of the hazard caused by fuel transportation by fuel trucks was estimated in a fragment of the Lithuanian roadway network. In this study non-constant immunity is analyzed.

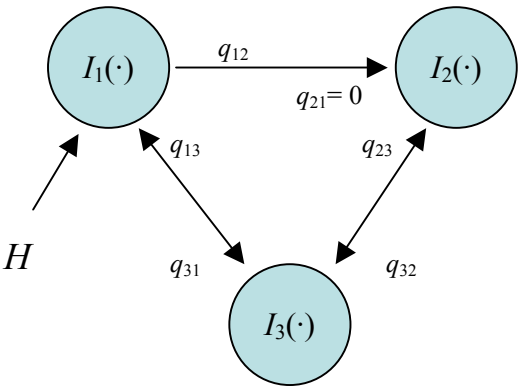

Figure 2: An example of a network system.

3 Node immunity function updating by application of the Bayesian approach

Commonly the Bayesian approach is applied to update estimated parameters of a stationary process when more statistical information becomes available. But often there is a need to deal with problems that are related to non-stationary processes. In such cases available statistical data cannot be used to update the characteristics of the previous period, because it represents the other state of the system. The required information for analysing a non-stationary process is

- distribution of statistical data;
- form of the trend of system dynamic describing characteristics (as functions of some factors and parameters), for example, it is exponential, polynomial, linear, etc.

In the presented task, the immunity of the chosen node depends on l different factors $F_1, \ldots, F_l$, and the trend of immunity is a priori known, so the expected value of immunity satisfies this equality

$$\mathrm{E}I(\theta, F_1, ..., F_k) = f(\theta, F_1, ..., F_k), \tag{3}$$

here $\theta = (\theta_1, \ldots, \theta_k)^{\mathrm{T}}$ the multidimensional parameter. Prior information can lead to some uncertainty and these parameters are assumed as random independent variables with their prior probability distributions.

Assume that parameters $\theta_1, \ldots, \theta_k$, density functions of prior distributions $\mathrm{pi}(x_i)$, $i = 1, \ldots, k$, distribution of statistical data I_j, $j = 1, \ldots, m$, (immunity in the j^{th} cycle) is also known, i.e. likelihood function $L(\cdot)$ that satisfies eqn (3). Posterior multidimensional density function is obtained by application of the Bayesian formula for this information [2]

$$\varphi(x_1,...,x_k \mid I_1,...,I_j) = \\ = \frac{\prod_{i=1}^{k} p_i(x_i) \cdot L(I_1,...,I_j \mid x_1,...,x_k)}{\int_{R_1}...\int_{R_k} \prod_{i=1}^{k} p_i(x_i) \cdot L(I_1,...,I_j \mid x_1,...,x_k) dx_1...dx_k}, \quad (4)$$

$j = 2, \ldots, m$, R_i is the range (set of all possible values) of parameter θ_i.

So the Bayesian estimate (expected value of posterior distribution) of parameter θ_i is

$$\hat{\theta}_{ij} = \int_{R_1}...\int_{R_i}...\int_{R_k} x_i \cdot \varphi(x_1,...,x_i,...,x_k \mid I_1,...,I_j) dx_1...dx_i...dx_k, \quad (5)$$

$j = 2, \ldots, m$.

3.1 Numerical experiment

Let's assume, that in the chosen node immunity has exponential dependence on the number of cycles i, i.e. the mean of the immunity is equal to its a priori known dependence function

$$\mathrm{E}I(a,i) = 1 - e^{-a \cdot i},\ a > 0. \quad (6)$$

Statistical data I_i: $0 \le I_i \le 1$, $i = 1,2, \ldots, m$, see fig. 3, were simulated by Beta distribution with the parameters

$$\alpha(a,i) = \sqrt{\frac{1 - e^{-a \cdot i}}{e^{-a \cdot i}}},\ \beta(a,i) = \frac{1}{\alpha(i)}. \quad (7)$$

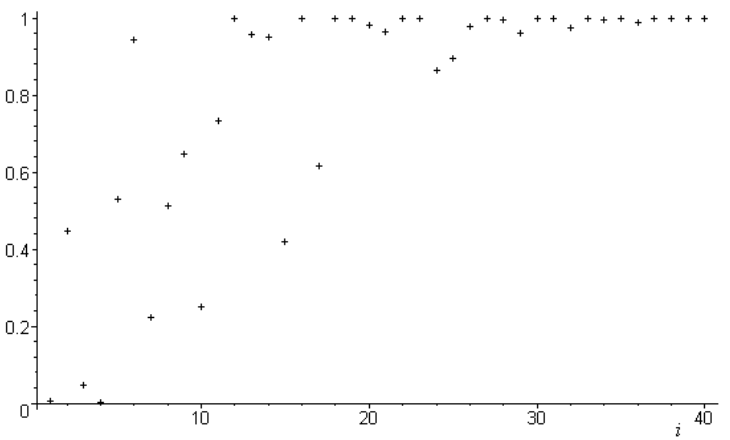

Figure 3: Simulated statistical data.

The value of parameter a was chosen as $a^* = 0{,}12$ for this simulation. The parameters of the Beta distribution were obtained considering requirement (3). Parameter β is expressed by α (eqn (7)) in order that expected value and variation of immunity have such properties

$$\mathrm{E}I(a,i) \underset{i\to\infty}{\to} 1,\ \mathrm{Var}I(a,i) \underset{i\to\infty}{\to} 0. \tag{8}$$

In this case, the likelihood function is

$$L(x,I_i) = \frac{1}{B\big(\alpha(x,i),\beta(x,i)\big)} I_i^{\alpha(x,i)}(1-I_i)^{\beta(x,i)}, \tag{9}$$

$i = 1,2, \ldots, m$, $B(\cdot)$ is the Beta function.

In the mathematical model a prior distribution of unknown parameter a is assumed as a non-informative distribution, i.e. the density function is

$$p(x) = const,\ x > 0. \tag{10}$$

In the case of vague prior knowledge and a large amount of data being available for updating a prior distribution, the usage of so-called non-informative prior distribution is useful, i.e. the posterior distribution is based on the information of likelihood function [3, 4].

The posterior distribution of parameter a is obtained by formula (4) using its prior density function and simulated data. So the density function of the posterior distribution is

$$p(x \mid I_1,\ldots,I_j) = \frac{p(x)\cdot\prod_{i=1}^{j} L(x,I_i)}{\int_0^\infty p(x)\cdot\prod_{i=1}^{j} L(x,I_i)dx} =$$

$$= \frac{p(x)\cdot\prod_{i=1}^{j} \frac{1}{B\big(\alpha(x,i),\beta(x,i)\big)} I_i^{\alpha(x,i)}(1-I_i)^{\beta(x,i)}}{\int_0^\infty p(x)\cdot\prod_{i=1}^{j} \frac{1}{B\big(\alpha(x,i),\beta(x,i)\big)} I_i^{\alpha(x,i)}(1-I_i)^{\beta(x,i)}dx}, \tag{11}$$

$j = 2, \ldots, m$, and Bayesian estimate (expected value) of parameter a is

$$\hat{a}_j = \int_0^\infty x\cdot p(x \mid I_1,\ldots,I_j)dx, \tag{12}$$

$j = 2, \ldots, m$.

The results (the estimates of the unknown parameter using the Bayesian approach) of the considered numerical experiment are shown in fig. 4. In this case the total sum of error squares is

$$\Delta = \sum_{j=1}^{m} (\hat{a}_j - a^*)^2 = 0{,}0329. \tag{13}$$

Updated immunity functions are presented in fig. 5.

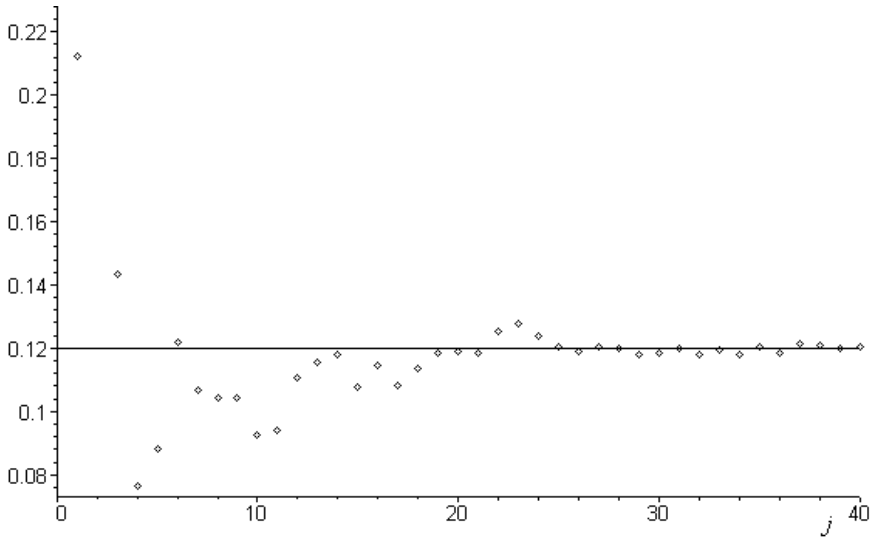

Figure 4: ◇– Bayesian estimates of random parameter a, —— – true value of parameter a (a^* = 0,12), j – number of Bayesian approach applications.

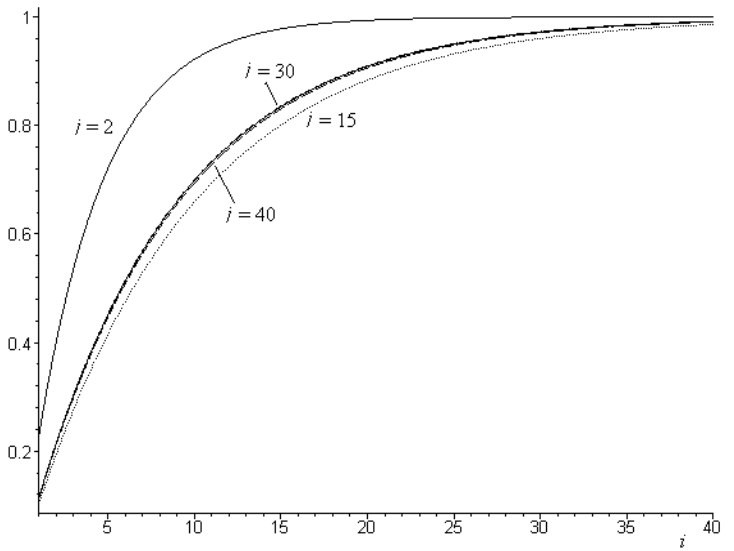

Figure 5: Graphics of true immunity function (——) and updated ones after the j^{th} Bayesian approach iteration, i – number of the cycle.

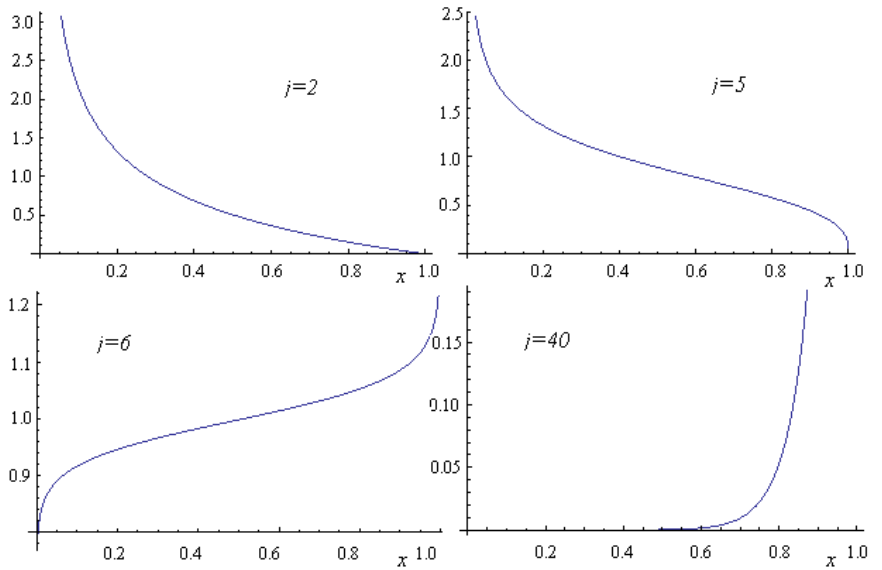

Figure 6: Transformation of statistical data (immunity) distribution density function, j – number of Bayesian approach applications.

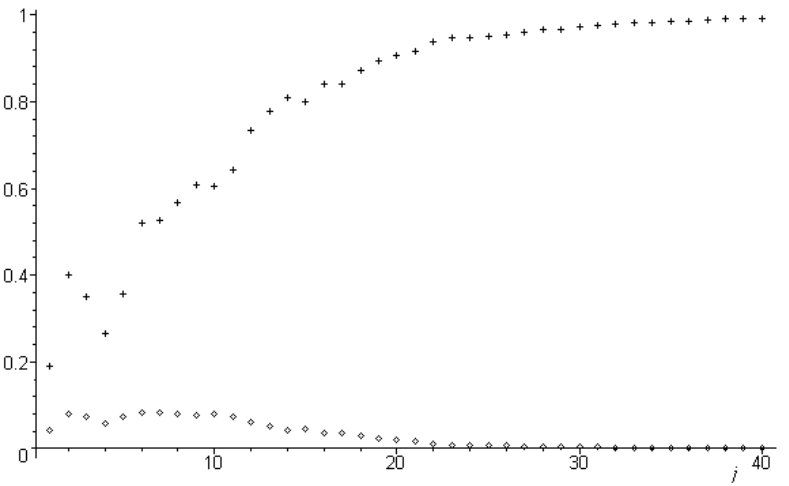

Figure 7: Means (+) and variances (⋄) of the distribution of node immunity, j – number of the cycle.

The transformation of statistical data (immunity) distribution density function obtained by the Bayesian approach application for updating random parameter a estimate is

$$g(x,\hat{a}_j)=\frac{1}{B\left(\alpha(\hat{a}_j,j),\beta(\hat{a}_j,j)\right)}x^{\alpha(\hat{a}_j,j)-1}(1-x)^{\beta(\hat{a}_j,j)-1}, \qquad (14)$$

$0 < x < 1$, j is the number of the cycle. Its graphs are presented in fig. 6, the means and variances of this distribution are presented in fig. 7.

4 Results and conclusions

The main aim of this paper is to present the developed mathematical model for network node immunity changes updated by the Bayesian approach and new available observations. With use of the presented algorithm the node immunity forecast is more and more precise (i.e., convergence to the true value); it gives the possibility to perform modelling of hazard distribution in the network with non-constant node immunities.

In the paper illustration of the developed model applicability is presented by numerical experiment. It was shown that the Bayesian approach application gives descending uncertainty of immunity describing distribution.

References

[1] Augutis, J., Krikstolaitis, R., Urbonas, R. & Uščpuras, E. Hazard distribution and risk assessment in the network systems. *Stochastic environmental research and risk assessment*, Vol. 21, No. 1, pp. 51–61, 2006.

[2] Zutautaitė, I., Augutis, J., Uspuras, E. Parameters estimation in ageing models using Bayesian approach. *Modeling of complex systems and environments: proceedings of the ISSAT international conference*, pp. 173–177, 2007.

[3] Bernardo, J. M., Smith, A. F. M. *Bayesian theory*. John Wiley & Sons, pp. 357–361, 2003.

[4] Berthold, M., Hand, D.J. *Intelligent Data Analysis*. 2nd edition, pp. 138–139, 2003.