On the Advanced Methodology of Risk-Based System Resilience Analysis

Yuriy V. Kostyuchenko
Scientific Centre for Aerospace Research of the Earth, National Academy of Sciences of Ukraine, Kiev, Ukraine.
E-mail: yuriy.v.kostyuchenko@gmail.com

(Received February 7, 2020; Accepted April 30, 2020)

Abstract
The modern evolution of technological systems from Hierarchical branching structures purposed to centralized transfer and distribution of limited resources to multi-agent interconnected self-organized networks aimed to produce, transport and consumption of resources are considered. The model of multi-agent interconnected self-organized adaptive networking systems is proposed, the network topology is considered, a system functioning model including transient processes is analyzed. A substantial limitation of the traditional reliability paradigm for a novel type of systems is demonstrated. It was assumed, that optimization approaches in the context of “big data” utilization lead to create a quasi-infinite space of non-structured decisions, which can be characterized as “big decisions”. The modified approach based on the “equally defended networked system” paradigm and the corresponding quantitative risk measure is proposed.

Keywords- System reliability, System structure, Topology, Big data, Big decisions, Risk, Equally-defended systems.

1. Introduction
The task of system resilience, analysis of generated by systems failures risks, is an important part of applied mathematics and engineering. Many decades this issue investigating from many sides: optimal decision making, risk and uncertainty analysis, system’s design and architecture optimization, management and control methods advancement, etc., with a wide range of research methods (Singh and Billinton, 1977; Ram and Singh, 2009; Kadry and El Hami, 2015; Kumar et al., 2017).

Traditionally, the task of reliability control may be considered in the framework of a well-known paradigm, as the difference between the expected and the real system performances. Usually, the real and targeted infrastructure functions at time $t$ are defined as $F(t)$ and $F_0(t)$. If $R$ denotes resilience, its value can be expressed as (Natvig, 2010; Zio, 2013):

$$R = \int_t F_0(t) \frac{F(t)-F(t)}{F_0(t)} dt,$$

(1)

With quantitatively metrics, included lots of linear and non-linear parameters, such as resilience capabilities, time to recover and other, for example (Høyland and Rausand, 2009; Kadry and El Hami, 2015):

$$r_i(S_p,F_r,F_d,F_0) = S_p \frac{F_r F_d}{F_0},$$

(2)
\[ S_p = \begin{cases} \frac{t_r}{t_r} e^{-\alpha(t_r-t_r^*)} & \text{for } t_r \geq t_r^* \\ \frac{t_r}{t_r} & \text{otherwise} \end{cases}, \] (3)

where, \( S_p \) is the speed recovery factor, \( F_o \) is the original system performance level, \( F_d \) is the performance level immediately after the disruption, \( F_r \) is the performance at a new stable level after recovery efforts have been exhausted, \( t_o \) representing slack time or the maximum amount of time post-disaster that is acceptable before recovery ensues, \( t_r \) is the time to final recovery or time to reach a new equilibrium state, \( t_r^* \) is the time to complete the initial recovery actions, and \( \alpha \) is the parameter controlling the “decay” in resilience until the new equilibrium is met (according to Hosseini et al., 2016).

Usually, this paradigm can be applied to the analysis of the wide class of engineering systems purposed to transition and distribution of centralized resources among the hierarchical network of consumers. In this case resilience analysis usually may be reduced to the task of optimization: minimization of distances between nodes and maximization of gain toward the links failure probabilities (for example, as \( \max \sum_j w_j F_{L,j}, with j \in L, and F \in M \)).

In these terms, the concept of resilience can be realized in different formalism according to the researcher’s priority.

But this paradigm developed to be applied first of all to engineering systems i.e. to the systems with specific structure and topology - hierarchical or branching structures with the one-direction distribution of limited resource.

Hierarchical branching structures can be described by the well-known replicator equation system, where dynamics of nodes \( N \) and links \( L \) may be represented as (Meisel and Gross, 2009; Peixoto, 2014):

\[ \frac{dN}{dt} = \begin{cases} N^2(1-N)(1-L), & \text{if } N \in [0,1] \\ 0, & \text{otherwise} \end{cases} \] \( (4) \)

\[ \frac{dL}{dt} = \alpha(N - N^*). \] \( (5) \)

In equation (31) the parameter \( \alpha \) controls the strength of feedback from the environment to the nodes. Correspondingly, in such representation, the errors will depend on entropy described as (Lee et al., 2011):

\[ S = -\sum_{r,s} e_{rs} \frac{e_{rs}}{2} - \sum_k N_k \ln(k!) - \frac{1}{2} \sum_{r,s} e_{rs} \ln \frac{e_{rs}}{e_r e_s}, \] \( (6) \)

Here \( N_k \) is the total number of nodes with degree \( k \), \( e_r = \sum_s e_{rs} \) is the number of the half-edge incident on block \( r \).

In such systems, the flows of data/resources and corresponding errors will be described as the linear combinations of parameters (in particular, as it presented by (2) and (3) equations).
The recent revolutionary development of technologies initiates new tendencies in the world, such as digitalization, globalization, and decentralization. The growth of these trends leads to the transformation now not solely of humans, but also of natural and engineering systems: all systems become multi-component, multi-physics, varied-scale, and distributed managing. All components (every node) of such a distributed system may be considered as a consumer, producer, or transmitter of resources and data depends on the current system’s topology, all agents are the objects and the subjects at the same time. In particular, an electrical network, based on “green energy” generation, instead of the traditional centralized power supply systems, can serve as an example of such a changed system. Therefore, the structure and function distribution of such open and expanded systems are changed, as well as the distribution of risks is changed too.

Development of improved paradigm of resilience, based on the new state of systems and changed the distribution of threats, may require a modified approach to maintain equal security of networks. This paradigm may be based on the known principle of equal strength design, which is actively used in the design of space technology (Tamaskar et al., 2014), and allows to manage a system’s complexity and reliability (Asikoglu and Simpson, 2012; Dolan and Lewis, 2008).

To identify the suitable paradigm of reliability and measures to control systems risks it is necessary to analyze the changed system’s structure, topology, and dynamics.

2. Methodological Notes
To describe a system structure and dynamics, let to estimate analyzed systems like the adaptive networks in terms of the distributed estimation problem. The way proposed is based on data diffusion normalization and data diffusion affine projection approach, which may be applied to systems with arbitrary distributions provided that the data statistical moments are known or may be forecasted (Lopes and Sayed, 2008).

2.1. General System Model: Multi-Agent Inter-Connected Self-Organized Adaptive Networking Systems
In terms of distributed estimation of the adaptive network, it should be considered the concept of diffusion adaptive networks (Lopes and Sayed, 2007; Lopes and Sayed, 2008) linked by the dynamic topology, in which the network nodes and corresponding links may be subject of failure.

To describe the data-normalized diffusion adaptive network, first of all, it is necessary to analyze Mx1 vector \(w^0\) from measurements collected at \(N\) network nodes (where each node \(k\) has a time realizations \(\{d_k(i), u_{k,i}\}, k = 1, ... N\) of zero-mean random data \(\{d_k, u_k\}\), with scalar measurement \(d_k(i)\) and a regression row vector \(u_{k,i}\).

After (Sayed, 2003) it may be assumed that measurements can be described by the linear model:

\[
d_k = u_k w^0 + v_k, \tag{7}
\]

Here \(v_k\) is a noise (with variance \(\sigma_{v,k}^2\)), which assumed spatially and temporarily independent.
We may assume here, that excepting the regressors \( u_{k,i} \), all vectors are column vectors (regressors are the row vectors). The random quantities will denote as \( \{d_k, u_k\} \), and its realizations as \( \{d_k(i), u_{k,i}\} \).

The adaptive network approach (Sayed, 2003; Lopes and Sayed, 2008) may be applied to estimate \( w^0 \). An adaptive network is a result of the application of local adaptive/learning rules or filters to nodes of the network. To efficiently exploit spatial and temporal information, the resulted communication topology exploiting the realization of a cooperation protocol between the nodes. Different cooperation protocols, based on the varied adaptive/learning rules or filters, generating the different adaptive networks.

Considering the adaptive network functioning under a diffusion protocol, each node \( k \) with neighborhood \( N_{k,i} \) at the time \( i \), should be analyzed, and the network should be defined as the set of nodes linked to \( k \).

In diffusion adaptive scheme every node \( l \in N_{k,i} \) has an estimate \( \psi_k^{(i-1)} \) with \( w^0 \); at each node \( k \) an aggregate estimate \( \phi_k^{(i-1)} \) through the linear combining the neighbors’ estimates is generating, and then the local estimate is updated from \( \psi_k^{(i-1)} \) to \( \psi_k^{(1)} \), according to (Sayed, 2003; Lopes and Sayed, 2008):

\[
\phi_k^{(i-1)} = \sum_{l \in N_{k,i}} c_{k,l}(i) \psi_l^{(i-1)}, \tag{8}
\]

\[
\psi_k^{(i)} = \phi_k^{(i-1)} + H_k u_{k,i}(d_k(i)) - u_{k,i} \phi_k^{(i-1)}, \tag{9}
\]

here the coefficients \( \{c_{k,l}\} \) are the set of local combiners (with \( \sum_l c_{k,l}(i) = 1 \); \( H_k \) is a matrix of local neighborhood regressions (Lopes and Sayed, 2007).

In (8) the aggregation mapping \( \psi_k^{(i-1)} \) could be presented by any broad (usually nonlinear) function of the close estimates: Metropolis-Hastings, closest neighbor, the Laplacian algorithm, or Bayesian rules.

Calculation (8)-(9) consolidates the effect of a few adaptive filters, as well as a dynamically changing topology of the network. To calculate the system performance the state-space formalization, with introducing global random quantities (according to Sayed, 2003; Lopes and Sayed, 2008) can be utilized:

\[
\psi^i \triangleq \text{col}\{\psi_1^{(i)}, ..., \psi_N^{(i)}\}, \tag{10}
\]

\[
\phi^i \triangleq \text{col}\{\phi_1^{(i-1)}, ..., \phi_N^{(i-1)}\}, \tag{11}
\]

\[
U^i \triangleq \text{diag}\{u_{1,i}, ..., u_{N,i}\}, \tag{12}
\]

\[
d_i \triangleq \text{col}\{d_i(i), ..., d_N(i)\}, \tag{13}
\]

\[
H(U^i) \triangleq \text{diag}\{H_1, ..., H_N\}, \tag{14}
\]
which gathers the information over all \( N \) nodes at a specific time.

A **global state-space network model** for (8)-(9) may be given following the (Sayed, 2003; Lopes and Sayed, 2008) as:

\[
\psi^i = G_i \psi^{(i-1)} + Hu^i(d^i - U_iG_i\psi^{(i-1)}),
\]

(15)

where \( G_i = C_i \otimes I_M \).

The presented model catches the dynamic system topology as the combining matrix \( C_i = [c_{k,l}(i)] \), and can combine few diverse adaptive algorithms, time-fluctuating systems and networks, systems subject to connection and hub failures, and time delays. The data about the momentary system topology is caught by the matrix \( C_i \) (if node \( k \) is connected to node \( l \) at the time \( i \), \( c_{k,l}(i) \neq 0 \); \( C_i q_N = q_N \) with \( q = col\{1, \ldots, 1\}, N \times 1 \)).

Let to consider a random network topology model. For it consider undirected graphs \( (c_{k,l}(i) = c_{l,k}(i)) \). Modeling of the topology dynamics will be based on the assumption of the links and nodes are random values (at the time \( i \), the random link \( c_{k,l}(i) \), connects nodes \( k \) and \( l \), has a nominal value \( c_{k,l}(i) = c_{l,k}(i) \) with probability \( p_{k,l} = p_{l,k} \), or will be zero with probability \( q_{k,l} \)) as (Sayed, 2003; Lopes and Sayed, 2008):

\[
[C_i]_{k,l} = \begin{cases} 
  c_{k,l} & \text{with } p_{k,l} \\
  0 & \text{with } q_{k,l} = 1 - p_{k,l}
\end{cases}
\]

(16)

The similar may be utilized to model the nodes: a probability of occurrence equals to \( p_{k,k} \), a nominal topology \( C_0 \), contained a fixed number of nodes \( N \) and \( n_l \) links, subjected by failures, the \( n_l \) links give \( 2^{n_l} \) different sub-networks \( C_i \) with probability \( p_l \) each, with existing and faulty links, so in such formalization the probabilities \( \{p_l\} \) are related to the \( \{p_{k,l}\} \).

In this model, the mean topology matrices \( G = E G_i \) and \( \bar{G} = E (G_i \otimes G_i^T) \) may be defined as:

\[
G = \Sigma_{l=1}^{n_l} p_l G_l,
\]

(17)

\[
\bar{G} = \Sigma_{l=1}^{n_l} p_l (G_i \otimes G_l^T),
\]

(18)

where \( p_l = P\{C_l = C_i\} \) and \( G_i = (C_i \otimes I_M) \).

### 2.2. System Functioning Model

To understand network resilience, transient processes should be analyzed.

Let suppose, that (Sayed, 2003; Lopes and Sayed, 2008):

\[
w^{(0)} = q_N \otimes w^0,
\]

(19)

\[
\tilde{\psi}^i = w^{(0)} - \psi^i.
\]

(20)
So, from the data model (7) can be obtained:

\[ d_i = U_{ki} w(0) + v_k, \]  

where \( v_i = \text{col}\{v_1(i), \ldots, v_N(i)\} \).

Because \( G_i w(0) = w(0) \), by subtracting \( w(0) \) from the left side and \( G_i w(0) \) from the right side of (9), it can be obtaining the following form:

\[ \tilde{y}_i = (I_{NM} - HU_i^* U_i) G_i \tilde{y}_i^{(i-1)} - HU_i^* v_i. \]  

Basing on the assumptions of (i) temporal and spatial independence of the regressors, and (ii) the collected data is not correlated with the network perturbations, the following form of expectations may be proposed (Lopes and Sayed, 2008):

\[ E[\tilde{y}_i] = [(I_{NM} - E(HU_i^* U_i))E G_i E \tilde{y}_i^{(i-1)}]. \]  

This form demonstrates, that the mean evolution of the vector of global weight error depends mainly on data moment \( E(HU_i^* U_i) \) and on mean topology matrix \( E G_i \).

In such a case, the global Mean-Square Deviation and Excess Mean-Square Error can be presented as (Lopes and Sayed, 2008):

\[ MSD = \frac{1}{N} E \| \tilde{y}_i^{(i-1)} \|^2, \]  

\[ MESE = \frac{1}{N} E \| \tilde{y}_i^{(i-1)} \|^2_{R_u}, \]  

where \( R_u = E U_i^* U_i \).

The mean-square network evolution model in terms of weighted norms of \( \tilde{y}_i^{(i-1)} \) and \( \| \tilde{y}_i^{(i-1)} \|^2 \Sigma \) can be calculated using energy conservation parameters (Sayed, 2003; Lopes and Sayed, 2008) (from equations (24) and (25) with \( \Sigma = I_{NM} \) or \( \Sigma = R_u \)).

At the node \( k \) the local output estimation error may be defined as:

\[ e_k(i) = d_k(i) - u_k \phi_k^{(i-1)}, \]  

And so the global error vector across the network \( e_i = \{e_1(i), \ldots, e_N(i)\} \) will be defined as:

\[ e_i = d_i - U_i \phi_i^{(i-1)} = U_i G_i \tilde{y}_i^{(i-1)} + v_i, \]  

and with including weight error from (15):

\[ \tilde{y}_i = G_i \tilde{y}_i^{(i-1)} - HU_i^* e_i. \]
From (28) through the procedure of energy balance, can be obtained expectation:

$$E\|\tilde{\gamma}^i\|^2 = E\|\tilde{\gamma}^{(i-1)}\|^2 + E\nu_i^T U_i^T \Sigma U_i^T \nu_i,$$

(29)

where:

$$\Sigma' = E G_i^T \Sigma G_i - E G_i^T \Sigma U_i G_i - E G_i^T U_i^T \Sigma H U_i^T G_i.$$

(30)

The recursive variance equation (28) describes the evolution of weighted norms of \( \tilde{\gamma}^i \) in terms of data statistical moments.

In equation (30) the weighting matrix \( \Sigma \) is implicit. To find expressions for the mean-square deviation and mean-square error it is necessary to use vectorization techniques (Sayed, 2003; Lopes and Sayed, 2008). Therefore, the particular solution may be obtained here for evolving topologies with Gaussian regressors in conditions with Gaussian transformation as follow:

$$H = D = diag\{\mu_1 I_M, ..., \mu_N I_M\}.$$  

(31)

$$\tilde{\gamma}^i = T^* \tilde{\gamma}^i,$$  

(32)

$$U_i = U_i T,$$  

(33)

$$\tilde{G}_i = T^* G_i T_i,$$  

(34)

Where: \( R_u = T \Lambda T^* \), \( T \) is unitary, and \( \Lambda = diag\{\Lambda_1, ..., \Lambda_N\}, \Lambda_k > 0 \) and diagonal.

Therefore, it can be obtained the next recursion, which describes the mean-square network performance in terms of parameter \( \tilde{\sigma} = bvec\{\Sigma\} \) as:

$$E\|\tilde{\gamma}^i\|^2 = E\|\tilde{\gamma}^{(i-1)}\|^2_{\tilde{\sigma}^2} + \tilde{\sigma} b^T,$$

(35)

$$\tilde{F} = \tilde{G} I_{NM^2} - (I_{NM} \otimes \Lambda D) - (\Lambda D \otimes I_{NM}) + (D \otimes D) \tilde{\sigma},$$

(36)

where \( \tilde{G} = E(\tilde{G}_i \otimes \tilde{G}_i^T), b = bvec\{R_v D^2 \Lambda\}, R_v = \Lambda_v \otimes I_M, \Lambda_v = diag\{\sigma_{v,1}^2, ..., \sigma_{v,N}^2\}. \)

The global mean-square deviation can be calculated by choosing the parameter \( \tilde{\sigma} = bvec\{I_{NM}\} \) in equation (35), and excess mean-square error – by selecting \( \tilde{\sigma} = bvec\{\Lambda\} \) respectively.

3. Discussion: on the Evolution of Reliability Control in Changing Systems

Because the systems (4)-(5) as well as its transition processes and links failure probabilities (19)-(36) and (6) are substantially different, the “traditional resilience paradigm” (1)-(3) has an essential limitation in application to multi-agent adaptive networking systems (Bie et al., 2017).

Formally, here we consider a change of the “resilience paradigm” induced the change of the system described: hierarchical or branching structures with the one-directional distribution of limited
resource vs. multi-agent interconnected self-organized adaptive networking systems with reproducing and redistribution of resources between nodes.

The presented approach to the model allows to make the few important conclusions.

On the one hand, as the theory has shown, in particular, results (16)-(18) and (24)-(36), links fail and so limiting of resources and data exchanges between the nodes does not degrade the network performance significantly (and so infrastructure functions) for a wide range of link probability $p$ (Lopes and Sayed, 2008; Losada et al., 2012). In other words, much fewer communication resources may be utilized to achieve a pre-defined performance level (targeted function) in diffusion protocols in the multi-agent inter-connected self-organized adaptive networking systems toward the traditional systems.

On the other hand, this result shows that classical system stability theory (1)-(3) is not suitable enough for describing distributed adaptive multi-agent systems (Gross and Sayama, 2009). When the number of nodes is large, the number of solutions that can be considered “optimal” will also be infinite. And choosing the separate one will be impossible. Therefore, the presence of big data (i.e., when $N \to \infty$ and $L \to \infty$) inevitably leads to the presence of “big decisions” (Kostyuchenko et al., 2020).

In such circumstances, it may be offered the concept of “equally defended networked system”, which can be expressed as:

$$\sum_{i,k} (p_{k,l} - \alpha_{i,k} d_k) \leq \text{var}\{v_i\}, \quad (37)$$

Where $\alpha_{i,k}$ is a managing parameter (in some cases and separate classes of tasks reducing to the traditional parameter controlling the “decay” in resilience and to parameter $\alpha$ controls the strength of systems feedback).

Variability of this value inside the network clusters can serve as the risk parameter. And its quantitative measure of distributed defense of any separate network cluster can serve as a risk-based reliability indicator:

$$\beta = -E\|\bar{\psi}^i\|_2^{-2} \frac{(1-\bar{\psi}^i)}{\sigma \text{var}\{v_i\}}. \quad (38)$$

It should be noted, that in the case of hierarchical networks equation (38) can be reduced to Bayesian form with Gaussian distribution (Straub, 2009), which links this approach to the traditional resilience paradigm.

In the case of the application of (37)-(38) concept, artificial intelligence algorithms can be used to manage the separate inter-connected clusters of adaptive networks.

4. Concluding Remarks

Technological changes of the last two decades, recognized as the digital revolution (Messner et al., 2019; Sachs et al., 2019), have led to global transformations characterized by drastic and deep social and technological crises (Jaarsveld et al., 2018). Modern world communities facing complex threats and challenges related to the development of important multi-scale tendencies of
globalization, decentralization, and technological transformation. Development of these tendencies now not solely generates new kinds of nexus, nonlinear interdependencies, and dangers such as systems or chain risks, however additionally limiting the applicability of traditional approaches to risk analysis (Kostyuchenko, 2018), resilience and vulnerability assessment, not solely natural, but also environmental, social and technological dangers (Ermoliev et al., 2012). Observing problems in the area of crisis control is linked with the transformation of the studied system (Oves et al., 2018).

The sustainable management of a new kind of transforming systems is primarily should be based on the utilization of big data, processed primarily by machine learning methods. But it should be noted, that the big data utilized with formal algorithms lead to the creation of the quasi-infinite area of non-structured decisions, which can be defined as “big decisions” (Kostyuchenko et al., 2020).

Therefore, it can be formulated a qualitatively new task: a task of the optimal solution choice in this “big decisions” space, which can be decided based on the new risk analysis and system resilience context (Stock and Seliger, 2016). Taking into account the modified nature of the systems and changed threats, the systems resilience assessment should be also adapted to new types of data, risks, uncertainties, and models, as well as to the new structure of the systems.

In this study, a change of the “resilience paradigm” induced the change of the system analyzed was described: hierarchical or branching structures with the one-direction distribution of limited resource vs. multi-agent inter-connected self-organized adaptive networking systems with reproducing and redistribution of resources between nodes. The modern evolution of technological systems is considered. The model of multi-agent inter-connected self-organized adaptive networking systems is proposed, the network topology is considered, a system functioning model including transient processes is analyzed. A substantial limitation of the traditional reliability paradigm for the novel type of systems has been demonstrated. It was assumed, that optimization approaches in the context of “big data” utilization lead to the creation a quasi-infinite space of non-structured decisions, which can be characterized as “big decisions”. The modified approach based on the “equally defended networked system” paradigm and corresponding quantitative risk measure is proposed.

Thus, the problem of creating a methodology for a decision-oriented method to assessing the system resilience based on an assessment of socio-environmental risk should be additionally investigated in further research. A formulation and approaches to solving the problem of "big solutions" can be proposed, based on the use of dynamic models. In particular, a robust solution may be proposed instead of optimal solutions; as well as the construction of the equally-defended distributed clustered system can be proposed as a way to improve the sustainability management of multi-agent inter-connected self-organized adaptive networking systems within the framework of integrated socio-environmental security management.

Conflict of Interest
The author confirms that there is no conflict of interest to declare for this publication.

Acknowledgements
The author is deeply grateful to the anonymous referees for constructive suggestions that resulted in important improvements to the paper, to colleagues from the American Statistical Association (ASA) and International Institute for Applied Systems Analysis (IIASA) for their critical and constructive comments and suggestions.
References


