Towards Reliability/Security Risk Metrics for Large-Scale Networked Infrastructures: Work in Progress

Vladimir Marbukh[a]
[a] NIST, Gaithersburg, USA, marbukh@nist.gov

Abstract: Economic and convenience benefits of interconnectivity drive the current explosive emergence and growth of networked systems. However, numerous systemic failures of various internetworked infrastructures demonstrate that interconnectivity also creates various risks, including risk of undesirable contagion. Our work in progress discusses challenges and possible approaches to developing reliability/security risk metrics for large-scale infrastructures, which quantify systemic risk of catastrophic phenomena. Since cascades leading to catastrophic phenomena in large-scale systems are possible due to cycles of positive feedbacks in system component interactions, conventional attack/fault tree models of multicomponent systems do not describe cascading phenomena. We propose to model component interactions by Markov field, which allow for such cycles to exist, and associate systemic failures with phase transitions as the number of system components becomes large. This model has advantage of benefiting from rich body of approaches and results provided by statistical physics. We carry out our analysis and interpretations under mean-field approximation which provides qualitative and sometimes even quantitative system description. We demonstrate a possibility of cascading behaviour leading to systemic failure. We argue that metrics of systemic risk in large-scale infrastructures should account for the likelihood of catastrophic transition within system time horizon. The risk of this transition becomes essential as system operational time horizon becomes comparable with “life expectancy” of the normal/operational metastable system equilibrium. We consider “large deviation regime” in which system operational time horizon is much less than the “life expectancy” of the normal/operational metastable system equilibrium, but high level of risk averseness makes systemic risk essential. Note phenomenological nature of our analysis of transitions between metastable states since consistent analysis should be based on the underlying Markov dynamics of the system. Such a consistent analysis as well as applicability of the proposed approach to real-life systems may be areas of future research.

1. INTRODUCTION

Economic and convenience benefits of interconnectivity drive the current explosive emergence and growth of networked systems. However, numerous systemic failures of various internetworked infrastructures demonstrate that interconnectivity also creates various risks, including risk of undesirable contagion [1]. Due to reliance on networked infrastructures, understanding and ability to manage the fundamental risk/benefit trade-offs of interconnectivity is one of the most important challenges faced by modern society. In a case of undesirable contagion, e.g., due to propagating computer viruses, cascading failures or overload, the goal of system designers and operators is keeping system inside of the contagion-free region in space of system parameters. Since typically economic and competitive incentives drive system design and operation towards the boundary of this region, the nature of the contagion emergence, e.g., continuous or discontinuous, is of critical importance due to the occasional breach of this boundary caused by unavoidable uncertainties.

Our work in progress discusses challenges and possible approaches to developing reliability/security risk metrics for large-scale infrastructures, which quantify systemic risk of undesirable cascades leading to systemic failure. Since cascades in large-scale systems are possible due to cycles of positive feedbacks in system component interactions, conventional attack/fault tree models of multicomponent systems do not describe cascading phenomena, we propose to model component interactions by Markov field, which allows for such cycles to exist, and associate systemic failures with phase transitions as the number of system components becomes large. This model has advantage of benefiting from rich body of approaches and results provided by statistical physics. We carry out our analysis and interpretations under mean-field approximation which provides qualitative and sometimes even quantitative system description. We demonstrate a possibility of catastrophic phenomena leading to systemic failure. We argue that metrics of systemic risk in large-scale infrastructures should account for the likelihood of catastrophic transition within system time horizon. The risk of this transition becomes essential if the probability of system operational time horizon exceeding the "life expectancy" of the normal/operational metastable system equilibrium is comparable or exceeds the risk tolerance level of the system. We consider "large deviation regime" in which system operational time horizon is much less than the "life expectancy" of the normal/operational metastable system equilibrium, but high level of risk averseness makes systemic risk essential. Note phenomenological nature of our analysis of transitions between metastable states since consistent analysis should be based on the underlying Markov dynamics of the system. Such a consistent analysis as well as applicability of the proposed approach to real-life systems may be areas of future research.

The paper is organized as follows. Section 2 introduces loss function of a multicomponent system, which generalizes notion of structural function for monotonic system [2]. For systems with component interactions without cycles it is possible to obtain loss distribution which is the basis for risk measures. This situation is illustrated on an example of a system security model described by a probabilistic Attack Graph. Section 2 proposes Markov field model of multi-component system, which is consistent with local component interactions and allows for cycles in these interactions to exist. Mean-field approximation for this model indicates a possibility of metastability and catastrophic cascades in large-scale infrastructures where component interactions contain positive feedback cycles. Section 3 discusses conventional notion of Value at Risk in context of systemic risk of catastrophic cascades in large-scale infrastructures. Finally, Section 4 briefly summarizes and outlines directions of future research.

2. MULTI-COMPONENT SYSTEM

2.1. System Losses

Consider system whose state is characterized by binary vector $\delta = (\delta_1,..,\delta_N) \in \{0,1\}^N$. In the context of reliability of a N - component system, vector δ characterizes reliability status of all components: $\delta_n = 0$ if component $n = 1,..,N$ is operational, and $\delta_n = 1$ if this component fails. In the context of security of a system with N potential vulnerabilities, vector δ characterizes which of the potential vulnerabilities have been exploited: $\delta_n = 0$ if vulnerability has not been exploited, and $\delta_n = 1$ otherwise. Note that due to causal relationships between component failures or vulnerability exploits, vector δ takes values in some subset of $\{0,1\}^N$.

We assume that system state δ can be mapped to system economic loss $L(\delta)$, where function $L(\delta)$ satisfies the following properties: (a) $L(0) = 0$, (b) function $L(\delta)$ is increasing, i.e., $L(\delta^1) \le L(\delta^2)$ if $\delta^1 \le \delta^2$, for any binary vectors $\delta^1 = (\delta_n^1) \in \{0,1\}^N$ and $\delta^2 = (\delta_n^2) \in \{0,1\}^N$, and (c) each component/vulnerability is relevant, i.e., for each $n = 1,..,N$ there exists vector $\delta_{-n} := (\delta_k, k \ne n)$, such that $L(0,\delta_{-n}) < L(1,\delta_{-n})$. Partial ordering of vectors is defined with respect to all vector components: $\delta^1 \le \delta^2 \Leftrightarrow (\delta_n^1 \le \delta_n^2, n = 1,..,N)$. These assumptions define class of structures which generalize class of monotonic structures [2] for which loss function $L(\delta)$ is binary: $L(\delta) = 0$ or $L(\delta) = L > 0$ for $\delta \in \Delta$. Such an example is considered in the next subsection.

Assuming that unconditional probability distribution of vector $\delta = (\delta_1,..,\delta_N)$, $P(\delta)$ is known, system reliability or security risk due to failed component or, respectively, successfully exploited vulnerabilities is fully characterized by the corresponding probability distribution of system losses $L(\delta)$:

$$P(L) = \sum\nolimits_{\delta\in\{0,1\},L(\delta)\le L} P(\delta) . \tag{1}$$

However, evaluation of unconditional distribution $P(\delta)$ is generally a difficult and still open problem. Indeed, unconditional distribution $P(\delta)$ incorporates both system structure and conditional probabilities of activation of individual vulnerabilities q_n , $n = 1,..,N$, given that the required prerequisites have been satisfied. In practice, conditional probabilities q_n are estimated from historical data, e.g., probabilities of successful exploits of individual vulnerabilities can be found in the National Vulnerability Database (NVD) and Common Vulnerability Scoring System (CVSS) scores [3].

We separate system architecture, to be described below in this subsection, and state of environment $\sigma = (\sigma_1,..,\sigma_N) \in \{0,1\}^N$ which characterizes environment "willingness" to activate different vulnerabilities if the required prerequisites are satisfied. The reason for this separation is that in adversarial setting, exploits may be associated with certain cost for the adversary who may choose not to exploit the corresponding vulnerability. Conventional attack/reliability model assigns conditional exploit probabilities $q_n = E[\sigma_n]$ and assumes that random variables σ_n are jointly statistically independent for $n = 1,..,N$:

$$Q(\sigma) = \prod\nolimits_{n=1}^{N} q_n^{\sigma_n} (1-q_n)^{1-\sigma_n} . \tag{2}$$

Generalization, which assume certain correlations between likelihoods of activation different vulnerabilities, is straightforward.

It is common [4] to encode causal relationships between system failures/exploits by binary functions $\chi_n(\delta_{-n}) \in \{0,1\}$, $n = 1,..,N$, where $\chi_n(\delta_{-n}) = 1$ if prerequisites for successful exploit of vulnerability n are satisfied, and $\chi_n(\delta_{-n}) = 0$ otherwise. We assume functions $\chi_n(\delta_{-n})$ to be increasing with respect to partial ordering of vectors δ_{-n} . For example, if prerequisite for exploitation of vulnerability n is successful activation of both vulnerabilities v_k and v_m , then $\chi_n(\delta_k,\delta_m) = \delta_k\delta_m$. If prerequisite for activation of vulnerability v_n is successful activation of at least one vulnerability v_k or v_m , then $\chi_n(\delta_k,\delta_m) = \delta_k + \delta_m - \delta_k\delta_m$. The following equations, which directly follow from definition of functions $\chi_n(\delta_{-n})$, are analytical representation of system component interdependencies:

$$\delta_n = \sigma_n \chi_n(\delta_{-n}) , \tag{3}$$

$n = 1,..,N$. We view (3) as a system of N equations with respect to vector δ , given vector σ .

In cases when system (3) has unique solution, which include cases of multi-component systems whose component interactions do not include positive feedback cycles:

$$\delta_n = \sigma_n \varphi_n(\sigma_{-n}) , \tag{4}$$

mapping (4) allows for reformulation random system loss in terms of conditional distribution $Q(\delta)$ rather than unconditional distribution $P(\delta)$ [5]:

$$P(L) = \sum\nolimits_{\sigma\in\{0,1\},\mathrm{L}(\sigma)\le L} Q(\sigma) \ , \tag{5}$$

where renormalized loss function is

$$\mathrm{L}(\sigma) := L[\sigma_1\varphi_1(\sigma_{-1}),..,\sigma_N\varphi_N(\sigma_{-N})] . \tag{6}$$

In particular, (5)-(6) yields average loss

$$\bar{L} = E_{Q(\sigma)}[\mathrm{L}(\sigma)] . \tag{7}$$

However, for multi-component systems whose component interactions allowing positive feedback cycles, mapping (4) may not exist since state of environment $\sigma = (\sigma_1, .., \sigma_N) \in \{0,1\}^N$ is consistent with numerous sets of exploited vulnerabilities $\delta = (\delta_1, .., \delta_N) \in \{0,1\}^N$. This situation is considered in Section 3 of the paper.

2.2. Example: Probabilistic Attack Graph

Consider shown in Figure 2 popular toy example [4].

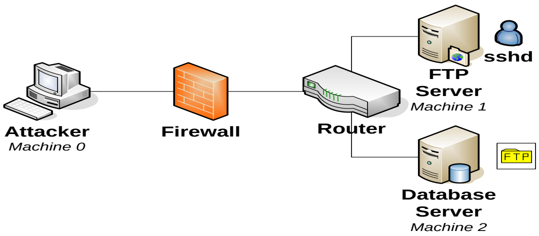


Fig. 1. Example of networked system

Machines 0, 1, and 2, are user's workstation, a web server, and a database server, respectively. The firewall allows http and ssh requests from machine 0 across to machine 1. During the normal operation, the user makes an http request to server 1, which goes through the firewall. Server 1 accesses database server running on server 2 to retrieve the required data and communicates back to machine 0 through http. If the user attempts to access machine 2 directly, e.g., by sending a ssh request from machine 0 to machine 2, the firewall blocks the communication. Successful attack may include a command injection attack on server 1 followed by a SQL injection attack on the database at machine 2. Then, the restricted data could be siphoned to server 1 and then to machine 0.

Attack graph for shown in Fig. 1 system is depicted in Fig. 2, where vulnerabilities are enumerated as follows: ftp_rhosts(0,1) $= v_1$, ftp_rhosts(0,2) $= v_2$, ftp_rhosts(1,2) $= v_3$, rsh(0,1) $= v_4$, rsh(0,2) $= v_5$, rsh(1,2) $= v_6$, sshd_bof(0,1) $= v_7$, local_bof(2) $= v_8$.

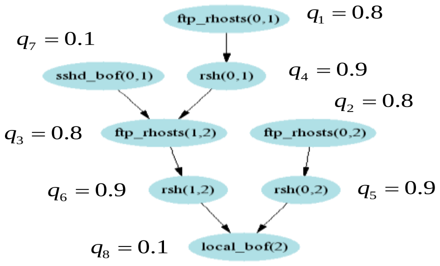


Fig. 2. Attack graph for shown in Fig. 1 system.

Following [4], we assume that point estimates of *conditional* probabilities of successful vulnerability exploits are as follows: $q_1 = q_2 = q_3 = 0.8$, $q_4 = q_5 = q_6 = 0.9$, $q_7 = q_8 = 0.1$. The corresponding functions $\varphi_n(\sigma_{-n})$ in (4) are as follows [5]: $\varphi_1 \equiv \varphi_2 \equiv \varphi_7 \equiv 1$, $\varphi_4(\sigma_{-4}) = \sigma_1$, $\varphi_5(\sigma_{-5}) = \sigma_2$, $\varphi_3(\sigma_{-3}) = \sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7$, $\varphi_6(\sigma_{-6}) = \sigma_3\varphi_3(\sigma_{-3}) = (\sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7)\sigma_3$, $\varphi_8(\sigma_{-8}) = \sigma_2\sigma_5 + \sigma_6\varphi_6(\sigma_{-6}) - \sigma_2\sigma_5\sigma_6\varphi_6(\sigma_{-6}) = \sigma_2\sigma_5 + (1-\sigma_2\sigma_5)(\sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7)\sigma_3\sigma_6$. Since economic loss L is due to user directly accessing machine 2, the renormalized loss function (6) is $\mathrm{L}(\sigma) := L\sigma_8\varphi_8(\sigma_{-8})$, i.e.,

$$\mathrm{L}(\sigma) = \sigma_8[\sigma_2\sigma_5 + (1-\sigma_2\sigma_5)(\sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7)\sigma_3\sigma_6]. \tag{8}$$

Substituting (8) into (5) yields probability distribution of system losses $P(L(\delta) = L) \approx 0.087$ and $P(L(\delta) = 0) = 1 - P(L(\delta) = L) \approx 0.013$ [5]. Thus, the expected loss (7) is $\overline{L} \approx 0.087L$.

3. MARKOV FIELD MODEL OF MULTI-COMPONENT SYSTEMS

3.1. Loss Distribution and Mean-Field Approximation

In a general case of multi-component system allowing feedback cycles in component interactions, system (3) may not have solution, and second, or may have multiple solutions with respect to vector δ . In the first case, “local interactions” (3) cannot be simultaneously for $n = 1,..,N$ realized in the system with any unconditional distribution $P(\sigma)$, and ink the second case, equations (3) for $n = 1,..,N$ are consistent with multiple unconditional distributions $P(\sigma)$. These possibilities can be reformulated in terms of the activation probabilities of individual vulnerabilities conditioned on the status of other vulnerabilities $p_n(\delta_n | \delta_{-n})$. Equations (3) are consistent with the following conditional probabilities:

$$p_n(\delta_n | \delta_{-n}) = \begin{cases} q_n \chi_n(\delta_{-n}) & if \quad \delta_n = 1 \\ 1 - q_n \chi_n(\delta_{-n}) & if \quad \delta_n = 0 \end{cases}, \tag{9}$$

which may or may not uniquely define unconditional probability distribution $P(\delta)$ for vector $\delta = (\delta_1,..,\delta_N)$.

Given provability distributions $p_n(\delta_n | \delta_{-n})$, we propose to define the unconditional probability distribution $P(\delta)$ for vector $\delta = (\delta_1,..,\delta_N)$ as follows:

$$P(\delta) = Z^{-1} \prod\nolimits_{n=1}^{N} p_n(\delta_n | \delta_{-n}), \tag{10}$$

where normalization constant, which is called partition function in statistical physics, is

$$Z = \sum\nolimits_{\delta \in \{0,1\}^N} \prod\nolimits_{n=1}^{N} p_n(\delta_n | \delta_{-n}). \tag{11}$$

Definition (10)-(11) assumes that that unconditional probability distribution $P(\delta)$ is a specific form of Markov random field [6], where “strengths” of local interactions are “consistent” with given provability distributions $p_n(\delta_n | \delta_{-n})$. Since Markov random fields allow for existence of cycles in the system component interdependencies, assumption (10)-(11) can be used for modelling systemic risk of catastrophic phenomena in large-scale systems as $N \to \infty$. In particular cases when system component interdependencies do not have cycles, definition (10)-(11) takes form of the Bayesian belief propagation network and the corresponding risk evaluation formalism is known under various names specifying their “Bayesian” nature, e.g., Bayesian Attack Graph [7].

To demonstrate flexibility of Markov random field model (10)-(11), below we consider the following model of local interactions

$$p_n(\delta_n | \delta_{-n}) = \begin{cases} q_n[\gamma_n + (1-\gamma_n)\chi_n(\delta_{-n})] & if \quad \delta_n = 1 \\ 1 - q_n[\gamma_n + (1-\gamma_n)\chi_n(\delta_{-n})] & if \quad \delta_n = 0 \end{cases}, \tag{12}$$

where parameters $0 \le \gamma_n \le 1$ characterize strength of the positive feedback cycles in the activation of system vulnerabilities. Case $\gamma_n = 0$, model (12) takes form of model (9) with the highest strength of this positive feedback, and another extreme case $\gamma_n = 1$, $n = 1,..,N$ corresponds to a mutually independent activations of all vulnerabilities. Model (12) enhances model (9) by allowing for “spontaneous” activation of vulnerability $n = 1,..,N$ with positive probability $\gamma_n q_n$ even when $\chi_n(\delta_{-n}) = 0$, i.e., the neighbouring vulnerabilities are not activated. When $\chi_n(\delta_{-n}) = 1$, i.e., the neighbouring vulnerabilities are not activated, vulnerability n is activated with higher probability q_n , $0 \le p_{n0} < q_n \le 1$. Thus distribution (10)-(12) can be used as reliability model of multicomponent system where different components can fail spontaneously and failure of “neighbouring” components increases likelihood of this failure. We demonstrate in the next subsection that existence of these positive feedback cycles creates a possibility of systemic failures. Also note that generalization allowing for differentiated feedback from different components is straightforward.

It is known from statistical physics [8] that evaluation of distribution (10)-(11) is typically unattainable due to computational intractability of the partition function (11). Statistical physics developed approximations which produce qualitatively and sometimes even quantitatively accurate results. As an example, consider mean-field approximation [8] which assumes that

$$P(\delta) \approx \tilde{P}(\delta, p) := \prod\nolimits_{n=1}^{N} [\, p_n^{\delta_n} (1-p_n)^{1-\delta_n} \,], \tag{13}$$

where $p_n := E_P[\delta_n]$. Averaging equations (3) over distribution $\tilde{P}(\delta)$ we obtain the following system of non-linear fixed-point equations

$$p_n = \phi_n(p_{-n}), \tag{14}$$

where vector $p_{-n} = (p_k, k \neq n)$ and functions

$$\phi_n(p_{-n}) = \sum\nolimits_{\delta_{-n} \in \{0,1\}^{N-1}} p_n(\delta_n | \delta_{-n}) \prod\nolimits_{k \neq n} [\, p_k^{\delta_k} (1-p_k)^{1-\delta_k} \,]. \tag{15}$$

Due to Brouwer fixed-point theorem [9], mean-field system (14)-(15) has at-least one solution. It can be shown that for sufficiently small unconditional probabilities of exploits q_n, $n = 1,..,N$, system (14)-(15) has unique stable solution $\tilde{p}_* = (\tilde{p}_{*n})$ which can be associated with normal/operational system equilibrium and the corresponding steady-state loss distribution (1) is approximated by

$$\tilde{F}_*(L) = \sum\nolimits_{\delta \in \{0,1\}^N, L(\delta) \le L} \prod\nolimits_{n=1}^{N} [\, \tilde{p}_{*n}^{\delta_n} (1-\tilde{p}_{*n})^{1-\delta_n} \,]. \tag{16}$$

For sufficiently high probabilities q_n, system (14)-(15), in addition to the normal/operational state, may have other stable solutions which can be interpreted as describing metastable, i.e., persistent, system states with high losses.

3.2. Systemic Failures in Large-Scale Infrastructures under Mean-Field Approximation

To demonstrate a possibility of systemic failure in large-scale infrastructures on an example of a homogeneous system with large number of cycles in the system component interactions. In this system, which is reminiscent of system considered in [10], all N components/vulnerabilities can be associated with nodes in a regular homogeneous graph where each node has the same degree $d \ge 1$, and model (12) of local interactions:

$$p_n(\delta_n | \delta_{-n}) = \begin{cases} q[\gamma + (1-\gamma)\chi_n(\delta_{-n})] & if \quad \delta_n = 1 \\ 1 - q[\gamma + (1-\gamma)\chi_n(\delta_{-n})] & if \quad \delta_n = 0 \end{cases}. \tag{17}$$

We assume that successful activation of a vulnerability n is possible only when $t \in \{0,1,..,d\}$ out of d neighboring to node n vulnerabilities have already been activated, i.e., $\chi_n(\delta_{-n} | t) = 1$ if $\sum\nolimits_{k \in I_n} \delta_k \ge t$, and $\chi_n(\delta_{-n} | t) = p_0$ otherwise, where $I_n \subset \{n_{k1},..,n_{kd}\} \subset \{1,..,N\} \setminus \{n\}$ is the set of neighboring nodes for node $n \in \{1,..,N\}$.

In this case, system (14)-(15) has a homogeneous solution $p_n = p$ which satisfies the following single fixed-point equation:

$$p = q\phi(p), \tag{18}$$

where function

$$\phi(p|t) = \gamma \sum_{i=0}^{t-1} \frac{d!}{i!(d-i)!} p^i (1-p)^{d-i} + (1-\gamma) \sum_{i=t}^{d} \frac{d!}{i!(d-i)!} p^i (1-p)^{d-i}. \tag{19}$$

Since $\phi(0) = \gamma$, parameter $\rho := \gamma q$ can be naturally interpreted as the exogenous load on the system.

Solution to equation (18)-(19) is shown in Fig. 1 for different set of system parameters (q, γ, t).

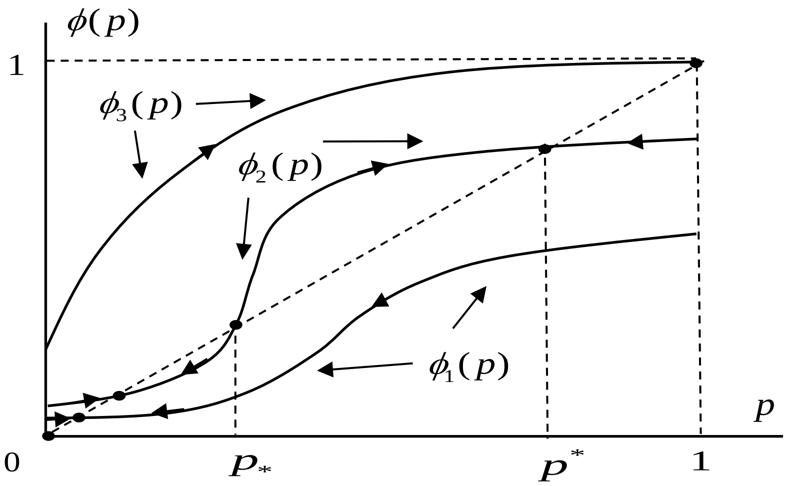
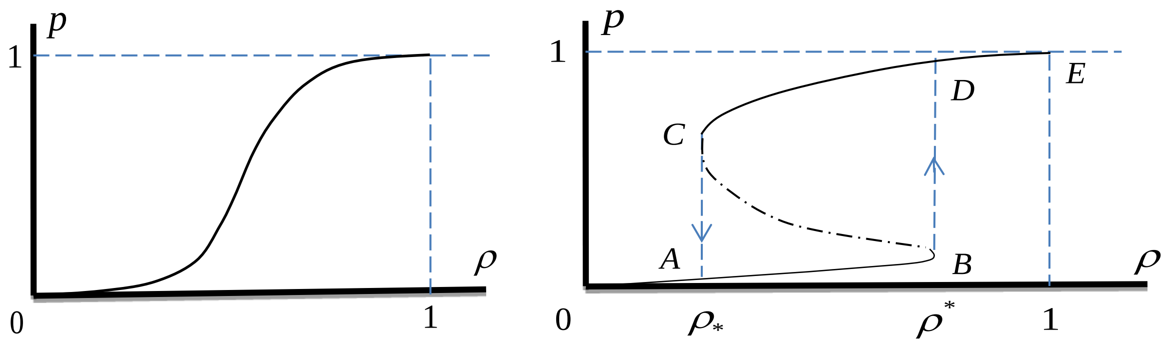


Fig. 3. Solution to fixed-point equation (18)-(19).

For sufficiently low exogenous load ρ and sufficiently low strength of the positive feedback of the vulnerability activations, e.g., characterized by sufficiently high parameter γ, function (19) is shown as $\phi(p) = \phi_1(p)$ in Fig. 3, and thus fixed-point equation (18)-(19) has unique globally stable solution $p = p_*$. For intermediate values of ρ and sufficiently high strength of the positive feedback of the vulnerability activations, e.g., characterized by sufficiently low parameter γ, function (19) is shown as $\phi(p) = \phi_2(p)$ in Fig. 3. In this case, fixed-point equation (18)-(19) in addition to stable equilibrium $p = p_*$ has another equilibrium $p = p^*$, $p^* > p_*$. For sufficiently high ρ, function (19) is shown as $\phi(p) = \phi_3(p)$ in Fig. 3, and thus fixed-point equation (18)-(19) has unique globally stable solution $p = p^*$. Following conventional interpretation of mean-field approximation, we interpret stable solutions p_* and p^* as describing normal/operational and catastrophic system equilibria respectively. Coexistence of these solutions as locally stable we interpret as describing metastable system equilibria.

Fig. 4 depicts solution to fixed-point equation (18)-(19) vs. exogenous load ρ in a case of low positive feedback, when this system has unique globally stable solution.



(a) **(b)**

Fig. 4a-b. Solution to fixed-point equation (18)-(19) vs. load ρ for low (**a**) and high (**b**) positive feedback.

Fig. 4 depicts solution to fixed-point equation (18)-(19) vs. exogenous load ρ in a case of high positive feedback, when this system may have multiple unique locally stable solutions.

For sufficiently low load $\rho < \rho_*$, system (18)-(19) has unique globally stable “good” solution $p = p_*$ represented by curve $0A$. For intermediate load $\rho_* < \rho < \rho^*$, this solution, represented by curve AB, is locally stable, coexists with locally stable “bad” solution $p = p^*$ represented by curve CD. For sufficiently heavy load $\rho > \rho^*$, this “bad” solution $p = p^*$, represented by curve DE, is unique and globally stable. In terminology of phase transitions [8], Fig. 5 indicates that system experiences discontinuous phase transition, also known as phase transition of first kind, which is associated with metastability and hysteresis loop $ABDCA$ as exogenous load ρ changes adiabatically, i.e., much slower than life expectancy of metastable states.

3. RISK METRICS FOR LARGE-SCALE NETWORKED INFRASTRUCTURES

3.2. Landau Theory Based Risk Metrics for Large-Scale Networked Infrastructures

Expected loss $\bar{L} = E_{P(\delta)}[L(\delta)]$ may not be an adequate representation of the security risk since $\bar{L}$ does not account for the tail risk. This motivated introduction of Value at Risk (VaR) [11]:

$$VaR_{1-\alpha} = \inf\{L \ge 0 : P(L(\delta) > L) \le \alpha\}, \tag{20}$$

where confidence level $1-\alpha$ quantifies decision maker risk averseness. Practical region for $VaR_{1-\alpha}$ lies between expected loss $\bar{L}$ for some $\alpha \approx 0.5$, and the maximum loss $\hat{L} := \max_{\delta\in\{0,1\}^N} L(\delta)$ for $\alpha = 0$. Further in the paper for simplicity we assume that system loss function is additive:

$$L(\delta) = \sum\nolimits_{n=1}^{N} l_n \delta_n , \tag{21}$$

where constants $l_n \ge 0$ characterize “importance” of vulnerability $n = 1,..,N$. Generalization is straightforward.

Consider the following family of distribution on $\delta \in \{0,1\}^N$, which depends on parameter $\lambda \in (-\infty,\infty)$:

$$\Omega(\delta;\lambda) = P(\delta)\exp[A(\lambda) - \lambda L(\delta)], \tag{22}$$

where

$$A(\lambda) = -\ln \sum\nolimits_{\delta\in\{0,1\}^N} \exp[-\lambda L(\delta)]P(\delta) . \tag{23}$$

and distribution $P(\delta)$ is given by (10)-(11). Note that family (22)-(23) includes $P(\delta)$ for $\lambda = 0$.

Consider function of $\lambda \in (-\infty,\infty)$, $\mathrm{A}(\lambda; L) = A(\lambda) - \lambda L$, where $L \ge 0$ is a fixed parameter. Due to function $A(\lambda)$ convexity for $\lambda \in (-\infty,\infty)$ and easily verified equality $\partial A(\lambda)/\partial\lambda\big|_{\lambda=0} = \bar{L} := E_{P(\delta)}[L(\delta)]$, function $\mathrm{A}(\lambda;\bar{L}) = A(\lambda) - \lambda\bar{L}$ is minimized over $\lambda \in (-\infty,\infty)$ for $\lambda = 0$ [ ]. Introduce “free energy” $F(L) = \mathrm{A}[\lambda(L); L]$, where $\lambda = \lambda(L)$ is unique solution to $\partial A(\lambda)/\partial\lambda = L$, and thus

$$\min\nolimits_{L\ge 0} F(L) = F(\bar{L}) = 0 . \tag{24}$$

It is known [12] that for large-scale systems with large number of vulnerabilities $N >> 1$, probability distribution of system losses has the following form:

$$P(L(\delta) > L) \sim e^{-F(L)}, \tag{25}$$

where free energy $F(L)$ is an extensive characteristic, i.e., is proportional to N, as $N \to \infty$, most likely values of system loss L are concentrated in close neighborhood of free energy $F(L)$ minima. Due to (24)-(25), system Value at Risk (20) can be approximated for $N >> 1$ as follows:

$$VaR_{1-\alpha} \approx \sup\{L \geq 0 : F(L) \leq -\ln\alpha\}\,. \tag{26}$$

Unfortunately, for systems with large number of interdependent vulnerabilities N, expression (26) cannot be used directly due to intractability of free energy $F(L)$.

Statistical physics developed various approximations for free energy $F(L)$. In the rest of this subsection we discuss a mean-field approximation $\tilde{F}(L) \approx F(L)$ in a situation when system (14)-(15) may have two stable solutions $\tilde{p}_* = (\tilde{p}_{*n})$ and $\tilde{p}^* = (\tilde{p}_n^*)$ describing normal/operational and catastrophic system equilibria with losses $\tilde{L}_*$ and $\tilde{L}^* >> \tilde{L}_*$ respectively. Fig. 6 shows a "relative" approximate free energy $\Delta\tilde{F}(\Delta L) := \tilde{F}(\tilde{L}_* + \Delta L) - \tilde{F}_*(\tilde{L}_*)$.

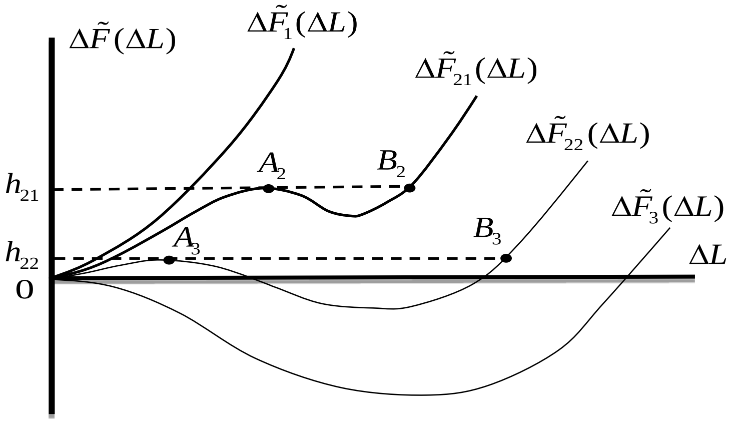


Fig. 5. System free energy under mean-field approximation.

Function $\Delta\tilde{F}_1(\Delta L)$ corresponds to a situation when mean-field system (14)-(15) has unique globally stable solution $\tilde{p}_* = (\tilde{p}_{*n})$ describing normal/operational system equilibrium. In Fig. 5 this situation corresponds to region $0 \leq \rho < \rho_*$. Functions $\Delta\tilde{F}_{21}(\Delta L)$ and $\Delta\tilde{F}_{22}(\Delta L)$ correspond to situations when mean-field system (14)-(15) has two locally stable solutions $\tilde{p}_* = (\tilde{p}_{*n})$ and $\tilde{p}^* = (\tilde{p}_n^*)$ describing normal/operational and catastrophic system equilibria respectively. In Fig. 5 these situations correspond to region $\rho_* < \rho < \rho^*$. Since vast steady-state probability that system loss L are concentrated in close proximity to the global minimum of function $\Delta\tilde{F}(\Delta L)$, functions $\Delta\tilde{F}_{21}(\Delta L)$ and $\Delta\tilde{F}_{22}(\Delta L)$ demonstrate situations when this steady-state system equilibrium is normal/operational and catastrophic respectively. In Fig. 5 these situations correspond to region $\rho = \rho_1$ and $\rho = \rho_2$ respectively, where $\rho_* < \rho_1 < \rho_2 < \rho^*$. Function $\Delta\tilde{F}_3(\Delta L)$ represents a situation of the stability boundary of solution $\tilde{p}_* = (\tilde{p}_{*n})$, which corresponds to point $\rho = \rho^*$ in Fig. 5.

3.2. Systemic Risk in Large-Scale Infrastructures under Large Deviation Regime

Presence of metastable states creates a possibility that observable random losses get stuck in some metastable equilibrium and do not have sufficient time to explore other metastable states or reach steady-state distribution. Assuming existence of two metastable states, normal/operational with low losses $L \approx L_*$ and catastrophic with unacceptably high losses $L \approx L^*$, and that initially, at moment $t = 0$, system resides at the normal/operational equilibrium, it is natural to consider the following time $t > 0$ dependent Value at Risk:

$$VaR_{1-\alpha}(t) = \inf\{L \geq 0 : P(t, L(\delta) > L) \leq \alpha\}\,. \tag{27}$$

Despite time evolution of $VaR_{1-\alpha}(t)$ depends on the underlying system dynamics, it is possible to obtain quantitative results in large deviation regime using with rate function derived from free energy [ ].

It is known from statistical thermodynamics that wide range of underlying system dynamics results in Langevin potential dynamics [12]. Analysis of this Langevin dynamics demonstrates [13]-[14] that lifetime of the normal/operational system state τ is distributed exponentially $P(\tau \leq t) \approx 1 - \exp(-t/\overline{\tau})$, where expected lifetime of this state $\overline{\tau} = E[\tau]$ is approximately exponential in N as $N \to \infty$:

$$\overline{\tau} \sim \theta e^{\Delta F^{+}(L)}\,, \tag{28}$$

where

$$\Delta F^{+}(\Delta L) = \int_{\bar{L}}^{L} [d\Delta F(\Delta L)/d\Delta L]^{+} dL\,, \tag{29}$$

parameter $\theta = O(1)$ as $N \to \infty$ characterizes time scale of macroscopic system relaxation and $[x]^{+} := \max(0, x)$. Function $F^{+}(L)$ in (29) is an extensive variable, i.e., $\Delta F^{+}(\Delta L) = O(N)$ as $N \to \infty$. This function represents local potential for the Langevin dynamics [12], and is derived from the corresponding action functional under large deviation regime [13].

In Fig. 6a-b curve $0ABC$ depicts free energy $\Delta F(\Delta L)$ and curve $0AB^{+}C^{+}$ depicts local potential $\Delta F^{+}(\Delta L)$ in cases of low and high positive feedback in vulnerability actualization.

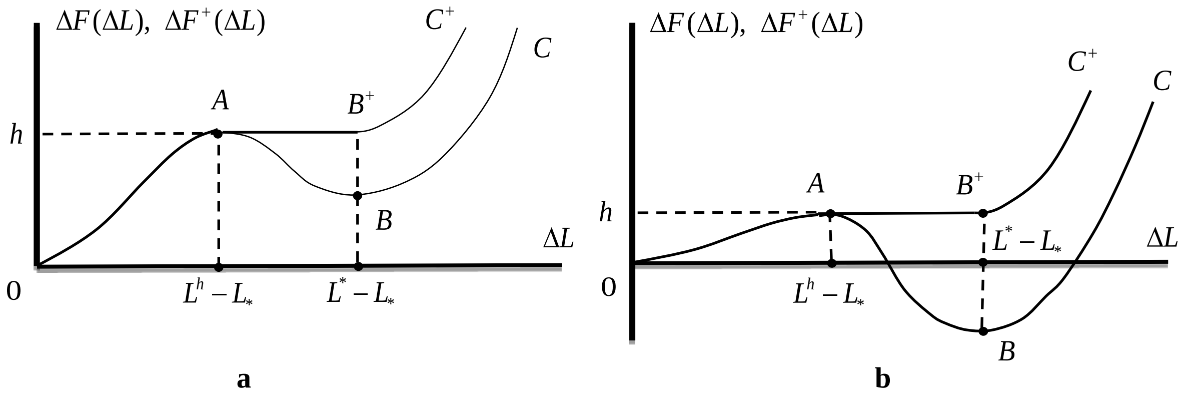


Fig. 6a-b. Free energy and local potential for low (**a**) and high (**b**) positive feedback.

Local minima of free energy $\Delta F(\Delta L)$ characterize system metastable states and global minimum characterizes system steady state. Fig 6a (6b) shows situation when normal/operational system equilibrium is system steady state (metastable) while catastrophic equilibrium is metastable (steady state). Local potential $\Delta F^{+}(\Delta L)$ quantifies “high of the barrier” required for the system to reach state with losses $L > L_{*}$ since point A in Fig. 6a-b represents the boundary of the “attraction region” of the normal/operational region.

We assume that system operational horizon T is much shorter than expected lifetime of the normal/operational system equilibrium $\overline{\tau}$: $T << \overline{\tau}$, and thus probability of transition to the catastrophic state during system lifetime

$$P(\tau \leq T) \approx 1 - \exp(-T/\overline{\tau}) \approx T/\overline{\tau} << 1\,. \tag{30}$$

We consider large deviation regime when both probability of system transition to catastrophic equilibrium during system lifetime $P(\tau \leq T)$ and tolerable risk α are small but comparable for

$N >> 1$, i.e., formally, $P(\tau \le T), \alpha \to 0$, but $\alpha^{-1}P(\tau \le T) = O(1)$ as $N \to \infty$.

Using (28) one can show that in this regime

$$VaR_{1-\alpha}(T) \approx L_* + \inf\{\Delta L \ge 0 : (T/\theta)e^{-\Delta F^+(\Delta L)} \le \alpha\}, \tag{31}$$

and thus

$$VaR_{1-\alpha}(T) \approx L_* + \Delta L, \tag{32}$$

where ΔL is an unique solution to equation

$$\Delta F^+(\Delta L) = \ln[T/(\alpha\theta)]. \tag{33}$$

Fig. 7 shows system time horizon Value at Risk (32)-(33) vs. "effective risk averseness" $\xi = T/(\alpha\theta)$.

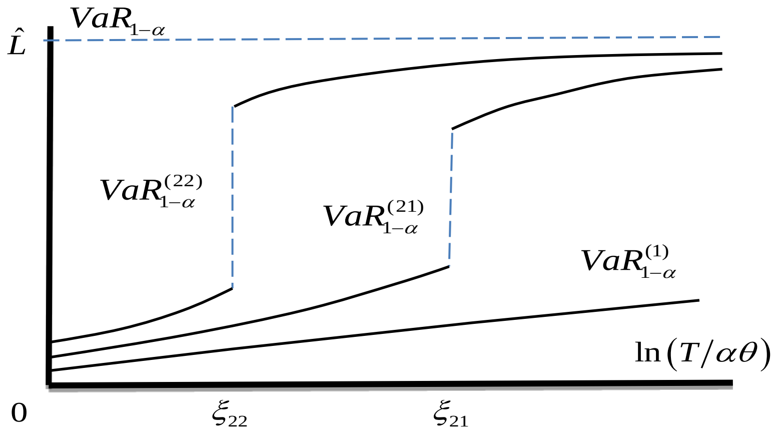


Fig. 7. System time horizon aware Value at Risk.

Curves $VaR^{(1)}_{1-\alpha}$, $VaR^{(21)}_{1-\alpha}$, and $VaR^{(22)}_{1-\alpha}$ in Fig. 7 correspond to local potentials $\Delta\tilde{F}_1(\Delta L)$, $\Delta\tilde{F}_{21}(\Delta L)$, and $\Delta\tilde{F}_{22}(\Delta L)$ in Fig. 5 respectively. Constants ξ_{21} and ξ_{22} are determined by conditions $\Delta\tilde{F}^+_{21}(L^h - L_*) = \ln \xi_{21}$ and $\Delta\tilde{F}^+_{22}(L^h - L_*) = \ln \xi_{22}$ respectively. Discontinuity in system time horizon aware Value at Risk indicate risk of system transition to catastrophic equilibrium through cascading process during system lifetime.

3. CONCLUSION AND FUTURE RESEARCH

This paper has reported on our work in progress on reliability/security risk metrics for large-scale infrastructures, which quantify systemic risk of undesirable cascades leading to systemic failure. Since cascades in large-scale systems are possible due to cycles of positive feedbacks in system component interactions, conventional attack/fault tree models of multicomponent systems do not describe cascading phenomena, we propose to model component interactions by Markov field, which allows for such cycles to exist, and associate systemic failures with phase transitions as number of system components increases. We argue that metrics of systemic risk in large-scale infrastructures should account for the likelihood of catastrophic transition within system time horizon. The risk of this transition becomes essential if the probability of system operational time horizon exceeding the "life expectancy" of the normal/operational metastable system equilibrium is comparable or exceeds the risk tolerance level of the system. We consider "large deviation regime" in which system operational time horizon is much less than the "life expectancy" of the normal/operational metastable system equilibrium, but high level of risk averseness makes systemic risk essential.

Numerous issues deserve further investigation. Evaluation of risk metrics for large-scale systems whose underlying dynamics is described by a Markov process with large number of locally interacting components may be possible with approach [15]-[16]. Large deviation regime implies system being sufficiently distanced from the point of phase transition to the catastrophic equilibrium. It is known [8] that mean-field approximation does not give qualitatively accurate system description in close proximity to a point of phase transition. Modern theories of phase transition [8] may provide an adequate apparatus for quantitatively accurate risk metrics for systems in close proximity to a point of phase transition. Extension of the proposed approach to more recent risk measures, e.g., Entropic Value at Risk (EVaR) [17], is straightforward. Moreover, in addition to advantages of EVaR as a risk measure, e.g., coherency, EVaR has advantage of applicability to large-scale systems due to natural connection to entropy maximization. The ultimate goal should be application to specific socio-technical systems experiencing abrupt transitions.

References

[1] D. Helbing, “*Globally networked risks and how to respond*”, Nature. 497, 51-59, 02 May 2013.

[2] R. Barlow and F. Proschan, Mathematical Theory of Reliability, Wiley, New York, 1965.

[3] CVSS “Common Vulnerability Scoring System (CVSS)”, Forum of Incident Response and Security Teams (FIRST), http://www.f irst.org/cvss/.

[4] L. Wang, T. Islam, T. Long, A. Singhal and S. Jajodia, “*An Attack Graph Based Probabilistic Security Metrics*”, 22nd IFIP WG 11.3 Working Conference on Data and Application Security, London, UK, July 2008.

[5] V. Marbukh, “*Towards robust security risk metrics for networked systems: work in progress*”, published and presented at IFIP/IEEE International Symposium on Integrated Network Management, May 17-21, 2021, virtual.

[6] R. Kindermann and J. Laurie Snell, “Markov random fields and their applications”, American Mathematical Society. Contemporary Mathematics, Vol. 1 (1980).

[7] N. Poolsappasit, R. Dewri and I. Ray, "*Dynamic security risk management using Bayesian Attack Graphs*", in IEEE Trans. on Dependable and Secure Computing, vol. 9, no. 1, pp. 61-74, Jan.-Feb. (2012).

[8] L.P. Kadanoff, "*More is the same; phase transitions and mean field theories*", J. of Statistical Physics. 137 (5–6), pp. 777–797, (2009).

[9] R. Kellogg, L. Bruce, Y. Tien-Yien, and A. James, "*A constructive proof of the Brouwer fixed point theorem and computational results*", SIAM Journal on Numerical Analysis. 13 (4): pp. 473–483, (1976).

[10] A. Majdandzic, B. Podobnik, S.V. Buldyrev, D.Y. Kenett, S. Havlin, and H.E. Stanley, “Spontaneous recovery in dynamical networks,” Nature Physics, Published Online, Dec. 2013.

[11] P. Jorion. Value at Risk: The New Benchmark for Managing Financial Risk (3rd ed.). McGraw-Hill (2006).

[12] U. Seifert. Stochastic thermodynamics, fluctuation theorems, and molecular machines, Reports on Progress in Physics, Volume 75, Number 12, November 2012.

[13] A. Shwartz and A. Adam. Large Deviations for Performance Analysis: Queues, Communications, and Computing. (1995).

[14] M.I Freidlin and A.D. Wentzell. Random Perturbations of Dynamical Systems. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] 260 (Second ed.). New York: Springer-Verlag (1998).

[15] V. Marbukh, “*Towards economically efficient mitigation of systemic risk of undesirable contagion*”, 31st European Safety and Reliability Conference (2021), virtual.

[16] V. Marbukh, “*Towards Landau Theory of Systemic Risk in Large-Scale Networked Systems: Work in Progress*”, NetSci 2022, Feb 8-11, (2022), virtual.

[17] S.A. Ahmadi-Javid, "*Entropic value-at-risk: A new coherent risk measure*", J. of Optimization Theory and Applications. 155 (3): 1105–1123, (2012).