Towards Robust Security Risk Metrics for Networked Systems: Work in Progress

Vladimir Marbukh
National Institute of Standards & Technology
100 Bureau Drive, Stop 8910
Gaithersburg, MD 20899-8910
marbukh@nist.gov

Abstract—**Security risk quantification is a necessary step in protecting critical resources in today's networked systems. Conventional security risk measures are based on the point estimates of the likelihoods of potential multi-step attacks that combine multiple vulnerabilities. Drawbacks of these measures are due to disregard for the tail risk, inherent inaccuracy of estimates of low probabilities, and reliance on the specific attacker(s) model. The recently proposed measure of cybersecurity risk - Cyber security Value at Risk (CyVaR), which is based on the VaR measure of financial risk, accounts for the tail risk. However, CyVaR still suffers from reliance on the specific attack model, and moreover has its own problems, e.g., it is not a coherent risk measure, which is currently considered to be a necessary trait of a risk measure. Following the recent trend of replacing VaR with the robust and coherent Entropic VaR (EVaR) as a financial risk measure, we suggest replacing CyVaR with CyEVaR. Using an example of a networked system and a highly motivated and capable attacker, we demonstrate that conventional risk measures may significantly underestimate the actual cybersecurity risk. Finally, we outline directions of future research.**

Keywords—Networked system, cybersecurity, vulnerabilities, risk metrics, Entropic Value at Risk.

I. INTRODUCTION

Real systems contain potential vulnerabilities which can be exploited by adversary(ies) in attempt to disrupt system operations. Emergence of large-scale critical mission infrastructures prone to systemic failures, makes quantitative evaluation of the overall system security risk not only exceedingly urgent but also very challenging due to the large number of potential vulnerabilities and non-linear contribution of individual vulnerabilities to the overall system risk. Conventional measurement of the security risk by the expected economic loss due to successful exploits of potential vulnerabilities is often inadequate due to its disregard for the tail risk, inherent inaccuracy of estimates of low probabilities, and reliance on the attacker(s) model. The recently proposed measure of cybersecurity risk - Cybersecurity Value at Risk (CyVaR) [1], which is based on the Value at Risk (VaR) measure of financial risk, accounts for the tail risk. However, CyVaR still suffers from low accuracy of VaR, and moreover has its own problems, e.g., is not a coherent risk measure, which is a highly desirable trait of an adequate risk measure.

Following the recent trend of replacing VaR with Entropic VaR (EVaR) as a measure of financial risk [2], we suggest replacing CyVaR with CyEVaR which is a robust and coherent risk measure. Robustness is achieved by accounting for attack model risk, i.e., allowing for controlled deviation of the attack strategy from the expected pattern. CyEVaR is defined as the worst-case expectation of the economic loss due to an attack within a feasible attack pattern. This definition, which has natural game-theoretic interpretation as a game against a boundedly rational adversary, interpolates between Bayesian and purely adversarial attack models. Coherency, i.e., subadditivity, which is currently considered to be a necessary trait of an adequate financial risk measure, is also highly desirable in cybersecurity since it encourages service diversification, e.g., multipath routing in networks or diversification in supply chains. Other desirable traits of CyEVaR include relative computational simplicity and sensitivity to low probability – high loss events which is important due to inaccuracy of estimates of low probabilities. Based on an example of a networked system with a highly motivated and capable attacker, we suggest that conventional risk measures may significantly underestimate the actual cybersecurity risk.

The paper is organized as follows. Section II introduces system/attacker(s) models, which quantify system loss due to successful exploits of potential system vulnerabilities and attacker(s) selection of potential vulnerabilities to exploit. Section III introduces and discusses Entropic Value at Cybersecurity Risk CyEVaR in context of other risk measures. Section IV illustrates CyEVaR measure on an example of a popular toy networked system whose cybersecurity is modelled by a probabilistic attack graph. Finally, Section V concludes and outlines directions of future research.

II. SYSTEM/ATTACK MODELS

Consider system with set of N potential vulnerabilities $\mathbf{V} = (v_n, n = 1,..,N)$ which can be exploited by adversary(ies). We identify set of actually exploited vulnerabilities with a random binary vector $\delta = (\delta_1,..,\delta_N)$, where $\delta_n = 1$ if vulnerability v_n is successfully exploited, and $\delta_n = 0$ otherwise. Due to causal relationships between component failures or vulnerability exploits, e.g., encoded by a fault tree, attack graph, or their combination [3], vector δ takes values in subset Δ of $\{0,1\}^N$: $\delta \in \Delta \subseteq \{0,1\}^N$. Set of

successful exploits $\delta = (\delta_1,..,\delta_N)$ causes system *economic loss* $L(\delta)$. We assume that (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 $\delta^1 = (\delta_n^1) \in \Delta$ and $\delta^2 = (\delta_n^2) \in \Delta$, and (c) each vulnerability is relevant, i.e., for each vulnerability v_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 [4] for which loss function $L(\delta)$ is binary: $L(\delta) = 0$ or $L(\delta) = L > 0$ for $\delta \in \Delta$.

Risk evaluation involves averaging over unconditional distribution of vector $\delta = (\delta_1,..,\delta_N)$, $P(\delta)$. However, evaluation of distribution $P(\delta)$ is generally a difficult and still open problem, especially for large-scale systems with large number of vulnerabilities. We 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 v_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 v_n is successful exploitations of both vulnerabilities v_k and v_m, then $\chi_n(\delta_k, \delta_m) = \delta_k \delta_m$. If prerequisite for exploitation of vulnerability v_n is successful exploitations of at least one vulnerability v_k or v_m, then $\chi_n(\delta_k, \delta_m) = \delta_k + \delta_m - \delta_k \delta_m$.

Unconditional distribution $P(\delta)$ incorporates both system structure and conditional probabilities of vulnerability exploits, given that the required prerequisites have been satisfied, and thus vulnerability v_n $n = 1,..,N$ can be in principle exploited. In practice, these conditional probabilities of $\tilde{q}_n$ are phenomenologically derived from the provided by the National Vulnerability Database (NVD) and Common Vulnerability Scoring System (CVSS) scores [5]. Our additional motivation is overcoming the limitations due to the assumption that attacker is oblivious to the system structure. We propose to achieve that by separating system and attacker(s) as follows.

While system is described by causal relationships between component failures or vulnerability exploits, attacker(s) strategy is described by random binary vector $\sigma = (\sigma_1,..,\sigma_N) \in \{0,1\}^N$, where conditional binary random variable $\sigma_n \in \{0,1\}$ assumes that $\chi_n(\delta_{-n}) = 1$, i.e., prerequisites for successful exploit of vulnerability v_n are satisfied. Given $\chi_n(\delta_{-n}) = 1$, component $\sigma_n = 1$ if vulnerability v_n is successfully exploited, and $\sigma_n = 0$ otherwise. Conventional attack/reliability model assigns conditional exploit probabilities $\tilde{q}_n = E[\sigma_n]$ and assumes that random variables σ_n are jointly statistically independent for $n = 1,..,N$:

$$\tilde{Q}(\sigma) = \prod_{n=1}^{N} \tilde{q}_n^{\sigma_n} (1 - \tilde{q}_n)^{1-\sigma_n} . \quad (1)$$

Viewing distribution (1) as a point estimate of the actual conditional distribution $Q(\sigma)$, robust risk measures account for possible inaccuracies in this estimate. Since distribution $Q(\sigma)$ quantifies the attacker(s) strategy, robustness implies a possibility that actual attacker(s) strategy deviates from the assumed one, which is a highly desirable trait in adversarial and highly uncertain cybersecurity decision making.

Starting point for our analysis are the following equations

$$\delta_n = \sigma_n \chi_n(\delta_{-n}), \quad (2)$$

$n = 1,..,N$, which directly follow from definition of functions $\chi_n(\delta_{-n})$. We view (2) as a system of N equations with respect to vector δ, given vector σ. Further we assume that this system has unique solution:

$$\delta_n = \sigma_n \varphi_n(\sigma_{-n}), \quad (3)$$

which is a case at least if the causal relations between different exploits do not have cycles. Mapping (3) allows for reformulation random system loss in terms of conditional distribution $Q(\sigma)$ rather unconditional distribution $P(\delta)$:

$$P(L(\delta) \le x) = Q(\mathrm{L}(\sigma) \le x), \quad (4)$$

where renormalized loss function

$$\mathrm{L}(\sigma) := L[\sigma_1 \varphi_1(\sigma_{-1}),..,\sigma_N \varphi_N(\sigma_{-N})] . \quad (5)$$

In particular, unconditional probabilities of exploits are

$$p_n = q_n E_{Q(\sigma)}[\varphi_n(\sigma_{-n})] . \quad (6)$$

III. Robust Security Risk Metrics

Expected loss

$$\tilde{L} := E_{\tilde{Q}}[\mathrm{L}(\sigma)] \quad (7)$$

may not be an adequate representation of the security risk since average (7) does not account for the tail risk. Even more importantly, point estimate of conditional probabilities $Q(\sigma) \approx \tilde{Q}(\sigma)$ may be highly unreliable.

Following Value at Risk (VaR) measure of financial risk, World Economic Forum has proposed notion of Cybersecurity VaR (CyVaR) [1]:

$$CyVaR_{1-\alpha} = \inf\{y \ge 0 : \tilde{Q}(\mathrm{L}(\sigma) \le y) \ge 1 - \alpha\}, \quad (8)$$

where confidence level $1 - \alpha$ quantifies decision maker risk averseness. Practical region for $CyVaR_{1-\alpha}$ lies between expected loss (7) for some $\alpha \in (0,1)$, and the maximum loss

$$\hat{L} := \max_{\sigma \in \{0,1\}^N} \mathrm{L}(\sigma) \quad (9)$$

for $\alpha = 0$. Serious deficiency of risk measure (8) is that $CyVaR_{1-\alpha}$ is not a coherent measure, i.e., violates subadditivity property [2], which is highly desirable since it encourages system redundancy and service diversification.

This and some other limitations of measure (8) motivated financial industry transition to Conditional Value at Risk (CVaR), also known as Expected Shortfall [6]. In our context, Cybersecurity CVaR (CyCVaR) takes the following form:

$$CyCVaR_{1-\alpha} = E_{\tilde{Q}}\left[\mathrm{L}(\sigma)\middle|\mathrm{L}(\sigma) > CyVaR_{1-\alpha}\right]. \quad (10)$$

While $CyCVaR_{1-\alpha}$ is a coherent risk measure, it inherits from $CyVaR_{1-\alpha}$ reliance on highly inaccurate estimate of the conditional probabilities $Q(\sigma) \approx \tilde{Q}(\sigma)$. This lack of robustness motivated financial industry to develop concept of robust risk measures. In our context, robust risk measures allow actual distribution $Q(\sigma)$ deviate from unreliable point estimate $Q(\sigma) \approx \tilde{Q}(\sigma)$. The rest of this paper suggests that an adequate cybersecurity risk measure could be based on Entropic Value at Risk (EVaR) which is a particular case of robust risk measures.

The corresponding Cybersecurity EVaR (CyEVaR) measure represents the maximum expected loss with respect to all feasible conditional distributions $Q(\sigma)$:

$$CyEVaR_{1-\alpha} = \max_{Q:H(Q\|\tilde{Q})\le -\ln\alpha} E_Q[\mathrm{L}(\sigma)], \quad (11)$$

where the Kulback-Leibler deviation is

$$H\left(Q\middle\|\tilde{Q}\right) := \sum_{\sigma\in\{0,1\}^N} Q(\sigma)\log\left[Q(\sigma)/\tilde{Q}(\sigma)\right] \quad (12)$$

and parameter $0 \le 1-\alpha \le 1$ characterizes decision maker risk eversness. It is easy to verify that $CyEVaR_{1-\alpha}$ increases from expected loss (7) to maximum loss (9) as $1-\alpha$ increases from 0 to 1. Briefly note that in addition to controlled robustness, risk measure (11) has other advantages, including relative computational tractability and natural game-theoretic interpretation, where boundedly rational adversary chooses probability distribution $Q(\sigma)$. Measure (11) interpolates between Bayesian model for $\alpha = 1$ and purely adversarial attack models for $\alpha = 0$.

In our context, power of adversary or attack severity is more naturally characterized by the expected, with respect to the conditional distribution $Q(\sigma)$, number of "conditionally successful exploits" $s_Q := \sum_{n=1}^N E_Q[\sigma_n]$, then parameter α. Risk measure (11), parameterized by $s = s_Q$ for $s > \tilde{s} := \sum_{n=1}^N \tilde{q}_n$, is given by solution to the following optimization problem

$$CyEVaR(s) = \max_{\alpha, Q:0\le\alpha\le 1, H(Q\|\tilde{Q})\le -\ln\alpha, s_Q\le s} E_Q[\mathrm{L}(\sigma)]. \quad (13)$$

Shown in Figure 1 function (13) increases from the expected loss (7) to the maximum loss (9) as s grows from $\tilde{s}$ to some $\hat{s} \le N$.

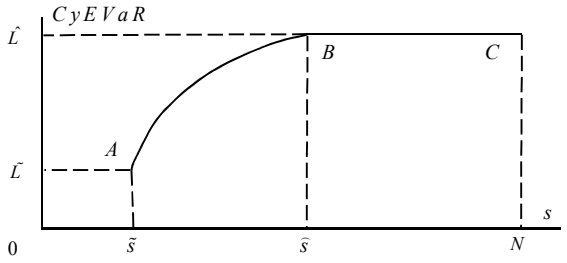

Figure 1. Entropic Value at Risk vs. attack severity.

Function $CyEVaR(s)$ is especially informative for large-scale systems with large number of potential vulnerabilities $N >> 1$ due to close connections of Entropic Value at Risk with Chernoff bound and large deviations [2]. In this short paper we only note that $CyEVaR(r)$ upper bounds the conditional expected loss, given attack severity:

$$E_{Q(\sigma)}\left[\mathrm{L}(\sigma)\middle|\sum_{n=1}^N \sigma_n \ge s\right] \le CyEVaR(s). \quad (14)$$

For system with large number of potential vulnerabilities N, under some technical conditions, inequality in (14) becomes equality asymptotically as $N \to \infty$. Also note that typically $\tilde{s} \le \hat{s} << N$ since our model assumes extremely determined and capable attacker to inflict system losses.

IV. Example: Probabilistic Attack Graph

Consider shown in Figure 2 popular toy example [7].

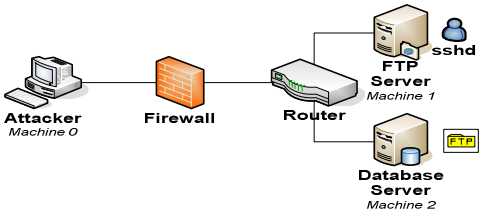

Figure 2. Example of networked system.

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 Figure 2 system is depicted in Figures 3, 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) $= v_8$.

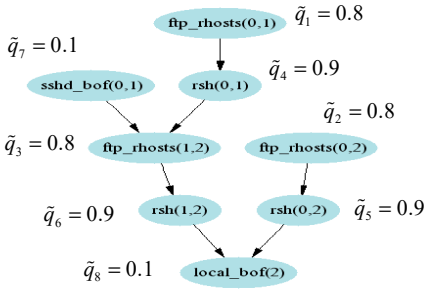

Figure 3. Attack graph for shown in Figure 2 system.

Following [7], we assume that point estimates of *conditional* probabilities of successful vulnerability exploits are as follows: $\tilde{q}_1 = \tilde{q}_2 = \tilde{q}_3 = 0.8$, $\tilde{q}_4 = \tilde{q}_5 = \tilde{q}_6 = 0.9$, $\tilde{q}_7 = \tilde{q}_8 = 0.1$.

The corresponding functions $\varphi_n(\sigma_{-n})$ in (3) are as follows:

$$\varphi_1 \equiv \varphi_2 \equiv \varphi_7 \equiv 1, \quad \varphi_4(\sigma_{-4}) = \sigma_1, \quad \varphi_5(\sigma_{-5}) = \sigma_2,$$

$$\varphi_3(\sigma_{-3}) = \sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7,$$

$$\varphi_6(\sigma_{-6}) = \sigma_3\varphi_3(\sigma_{-3}) = (\sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7)\sigma_3,$$

$$\varphi_8(\sigma_{-8}) = \sigma_2\sigma_5 + \sigma_6\varphi_6(\sigma_{-6}) - \sigma_2\sigma_5\sigma_6\varphi_6(\sigma_{-6}) =$$

$$= \sigma_2\sigma_5 + (1-\sigma_2\sigma_5)(\sigma_1\sigma_4 + \sigma_7 - \sigma_1\sigma_4\sigma_7)\sigma_3\sigma_6$$

and the renormalized loss function (5) is

$$\mathrm{L}(\sigma) := LE_{\tilde{Q}}[\sigma_8\varphi_8(\sigma_{-8})], \tag{15}$$

where economic loss due to user directly accessing machine 2 is L. In particular, the point estimates of unconditional probabilities of exploits $\tilde{p}_n = E_{\tilde{Q}}[\sigma_n\varphi_n(\sigma_{-n})]$ are as follows: $\tilde{p}_1 = \tilde{q}_1 = 0.8$, $\tilde{p}_2 = \tilde{q}_2 = 0.8$, $\tilde{p}_3 \approx 0.60$, $\tilde{p}_4 = \tilde{p}_5 = 0.72$, $\tilde{p}_6 \approx 0.54$, $\tilde{p}_7 = \tilde{q}_7 = 0.10$, $\tilde{p}_8 \approx 0.087$, and thus the expected loss is $\tilde{L} \approx 0.087L$.

It can be shown that solution to (13) is as follows:

$$Q(\sigma) = q(\sigma_2,\sigma_5,\sigma_8)\prod_{n=1,3,4,6,7}\tilde{q}_n^{\sigma_n}(1-\tilde{q}_n)^{1-\sigma_n}, \tag{16}$$

and thus

$$CyEVaR(s) = \max_{q(\sigma_2,\sigma_5,\sigma_8)} E_Q[(\tilde{p}_6 + (1-\tilde{p}_6)\sigma_2\sigma_5)\sigma_8], \tag{17}$$

where $\tilde{p}_6 \approx 0.54$ and maximization (17) is subject to the following constraints

$$\sum_{\delta_2,\delta_5,\delta_8\in\{0,1\}} q(\sigma_2,\sigma_5,\sigma_8)\log\left(\frac{q(\sigma_2,\sigma_5,\sigma_8)}{\prod_{n=2,5,8}\tilde{q}_n^{\sigma_n}(1-\tilde{q}_n)^{1-\sigma_n}}\right) \le -\ln\alpha \tag{18}$$

$$E_Q[\sigma_2 + \sigma_5 + \sigma_8] \le s - (\tilde{q}_1 + \tilde{q}_3 + \tilde{q}_4 + \tilde{q}_6 + \tilde{q}_7). \tag{19}$$

Solution (16)-(19) indicates that as attacker power, measured by the expected number of potential exploits s, increases from $\tilde{s} = \sum_{n=1}^{8}\tilde{q}_n = 5.4$ to $\hat{s} = \tilde{s} - (\tilde{q}_2 + \tilde{q}_5 + \tilde{q}_8) + 3 = 6.5$, the Entropic risk measure (17) grows from the expected loss $\tilde{L} = L\tilde{p}_8 \approx 0.087L$ to the maximum loss $\hat{L} = L$. Further increase in the attacker power does not increase the Entropic risk measure (17). This example demonstrates that in a case of highly determined and capable attacker, expected loss $\tilde{L}$ may significantly underestimate the actual cybersecurity risk. Further analysis indicates the same for measures (8) and (10).

V. Conclusion and Future Research

This paper suggests Entropic Value at Risk (EVaR) as a measure of cybersecurity risk. While EVaR has been gaining popularity as a measure of financial risk and has been extended to robust engineering of systems of various types, EVaR application to cybersecurity risk presents new opportunities and challenges. This paper, which is a work in progress, outlines some of them. Analysis of a simple networked system indicates that CyEVaR is a more adequate measure of cybersecurity risk than conventional measures at least in a case of highly determined and capable attacker. Our immediate plans include efficient evaluation of functions $\varphi_n(\sigma_{-n})$ in (3) for large-scale systems with a large number of potential vulnerabilities N and possibility of loops in the system Attack Graph. Imposing additional constraints on optimization (11)-(13) may allow for modelling attacker(s) of limited ability/determination. e.g., modelling limited cooperation of multiple attackers by imposing certain constraints on the mutual information of random variables σ_n, $n = 1,..,N$ in (11)-(13). Our ultimate goal is quantification of the existing qualitative cybersecurity risk mitigation recommendations, e.g., NIST Cybersecurity Framework [8]. Achieving this goal will require significant efforts in developing computationally effective yet accurate estimates of the Security Risk Reduction Return on Investment (SRR-RoI). In our future research we will employ computational techniques of statistical physics to approximate CyEVaR, and employ CyEVaR decomposition techniques, including Shapley value, to approximate SRR-RoI.

References

[1] World Economic Forum, "Partnering for Cyber Resilience Towards the Quantification of Cyber Threats," Report, 2015, available at http://www3.weforum.org/docs/WEFUSA_QuantificationofCyberThreats_Report2015.pdf

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

[3] R. Kumar and M. Stoelinga, "Quantitative security and safety analysis with attack-fault trees", 18th IEEE International Symposium on High Assurance Systems Engineering HASE, pp. 25-32, 2017.

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

[5] CVSS "Common Vulnerability Scoring System (CVSS)," Forum of Incident Response and Security Teams (FIRST), http://www.first.org/cvss/.

[6] R.T. Rockafellar and S Uryasev, "Optimization of conditional value-at-risk," Journal of Risk. 2 (3): 21–42, 2000.

[7] 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.

[8] NIST Cybersecurity Framework: https://www.nist.gov/cyberframework (links as of 26/5/20).