Scientific Research Publishing

Journal of Information Security, 2024, 15, 524-544
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

# Scale, Complexity, and Cybersecurity Risk Management

**Christopher Briscoe[1], Carl Young[1,2*]**

[1]Department of Mathematics and Physics, Emet Classical Academy, New York, NY, USA
[2]Consilience 360, New York, NY, USA
Email: *cyoung86@fordham.edu

**How to cite this paper:** Briscoe, C. and Young, C. (2024) Scale, Complexity, and Cybersecurity Risk Management. *Journal of Information Security*, **15**, 524-544.
https://doi.org/10.4236/jis.2024.154029

**Received:** June 21, 2024
**Accepted:** October 14, 2024
**Published:** October 17, 2024



## Abstract

Elementary information theory is used to model cybersecurity complexity, where the model assumes that security risk management is a binomial stochastic process. Complexity is shown to increase exponentially with the number of vulnerabilities in combination with security risk management entropy. However, vulnerabilities can be either local or non-local, where the former is confined to networked elements and the latter results from interactions between elements. Furthermore, interactions involve multiple methods of communication, where each method can contain vulnerabilities specific to that method. Importantly, the number of possible interactions scales quadratically with the number of elements in standard network topologies. Minimizing these interactions can significantly reduce the number of vulnerabilities and the accompanying complexity. Two network configurations that yield sub-quadratic and linear scaling relations are presented.

## Keywords

Complexity, Cybersecurity, Scale, Scaling Relations, Stochastic, Linear, Non-Linear, Macroscopic, Organized Complexity, Disorganized Complexity

## 1. Introduction

IT environments are variously described as complex, where complexity is frequently cited as a significant cybersecurity risk management issue. Yet despite its prevalence and impact, the precise effect on cybersecurity risk management is not well understood. Even cybersecurity tools purporting to reduce complexity are unable to articulate their effect except in very general terms.

In fact, fundamental questions about complexity remain unanswered, particularly with respect to cybersecurity. For example, what is an operationally useful

definition of complexity? What specific features of IT environments increase cybersecurity risk management complexity, *i.e.*, cybersecurity complexity, and by how much? What methods are effective in reducing cybersecurity complexity while preserving critical information management functionality?

The overall effect of cybersecurity complexity is to obfuscate cybersecurity risk management, which is defined as the application of security controls to vulnerabilities. However, describing its effect is necessary but not sufficient for developing an effective mitigation strategy.

In this paper we leverage an information theoretic model developed previously to examine the drivers of cybersecurity complexity mitigation pursuant to identifying effective mitigation on an enterprise scale [1]. Furthermore, we provide justification for a stochastic formulation of security risk management by defining complexity as a probabilistic phenomenon, which is an essential feature of the aforementioned model. Note that complexity in this context bears no relation to classical complexity theory, which studies the inherent difficulty of computational problems and the resources required to solve them.

Specifically, the information theoretic model in conjunction with a combinatorial analysis of specific network topologies yields quantitative relationships between the magnitude of cybersecurity complexity and the size of an IT environment. These relationships enable quantitative estimates of the effectiveness of specific mitigation strategies. Furthermore, security controls that address complexity *et al.* can be categorized according to their effect on the model parameters.

Addressing the fundamental questions noted above is the ultimate objective of this analysis. The most fundamental of these questions concerns the definition of complexity, and therefore is addressed first.

## 2. Complexity Described and Defined

A definition is the starting point for a rigorous treatment of any phenomenon since it ensures consistent use of terminology, provides the basis for subsequent analyses, and ultimately facilitates its integration into a theoretical framework. In this context, that framework is an information theoretic model of complexity, which itself is based on a simple stochastic model of security risk management.

Although the notion of complexity might seem intuitively obvious, intuition alone is insufficient to rigorously analyze its effects or the effectiveness of mitigation strategies. It is impossible to effectively analyze complexity and its impact on cybersecurity if it is unclear what complexity actually is.

Furthermore, although a truly general definition would apply to any complexity scenario, it is critical to be precise about the scenario in question. In this discussion the focus is on cybersecurity complexity, which specifically relates to cybersecurity risk management. Note complexity could equally relate to other issues such as IT administration. The complexity of one process could affect the other, but the features affecting the magnitude of each will be different in general.

Whatever else, it seems certain that complexity is not an *intensive* property. If

it were, its magnitude would be independent of size. A blue object consisting of 100 elements is just as blue as a similarly colored object with 1000 elements. In contrast, an object consisting of 1000 elements would typically be more complex than one with only 100 elements.

If complexity *only* depended on size, and each of the elements was independent, an entity consisting of N elements might always be twice as complex as an entity with N/2 elements. A property of an object that increases linearly with its size is an *extensive* property, by definition.

However, what if some or all of an object's constituent elements are identical? How would a lack of diversity affect the magnitude of complexity, if at all? For example, consider a jigsaw puzzle consisting of a thousand identical pieces. A blindfold is now placed on the puzzle solver, who is instructed to assemble the puzzle from scratch. Despite his visual impediment, the task is trivial because every possible puzzle configuration yields the correct result. In fact, the probability of guessing the correct configuration is one irrespective of the number of pieces. In other words, the puzzle assembly process has zero uncertainty, and complexity is not a function of system size.

Next consider a jigsaw puzzle consisting of 1000 *unique* pieces, which implies there is only one correct puzzle configuration. The probability that the blinded puzzle solver correctly positions the first piece is 1/1000. The probability he or she correctly positions the second piece is 1/999, etc. The probability of guessing the overall puzzle configuration is 1/1000! which is essentially zero. A similar calculation for puzzles containing even more pieces effectively yields the same result, which suggests that complexity increases exponentially with system size.

Note the blindfold enabled these simple calculations, where the probability of selecting the correct puzzle piece *at random* is just inversely proportional to the number of remaining pieces. If the blindfold were suddenly removed, the selection process would no longer be random since viewing the puzzle biases the configuration process. The additional puzzle information gleaned from viewing the puzzle would vitiate a calculation of the selection odds unless the bias itself were quantifiable. In short, the blindfold enables a straightforward application of the laws of probability.

In a similar vein, physicists achieved remarkable success using probability and statistics to describe certain macroscopic properties of materials. Recall such materials contain on the order of $10^{23}$ particles, which makes the application of Newton's Laws to each particle an intractable problem. The temptation is to attempt to replicate this success in other contexts by leveraging the benefits that accompany randomness. In particular, treating cybersecurity scenarios probabilistically would enable calculations identical to those used in the jigsaw puzzle example.

Clearly, the analogy between physical materials and IT environments has limits. The elements within IT environments do not behave erratically, and therefore the statistical methods that so effectively describe the average behavior of atoms and molecules aren't directly applicable. Moreover, problems in statistical mechanics

become manageable when particles are assumed to collide but don't otherwise interact. Arguably the opposite condition describes IT environments since networks facilitate interaction by design.

The applicability of the laws of probability hinges on the notion of randomness, which is also relevant to the concepts of organized and disorganized complexity. These ideas were posited in the late 1940s, and they reflected an attempt to distinguish between scenarios on the basis of the applicability of statistical physics.

Specifically, systems characterized by disorganized complexity exhibit the following conditions:

> "…the number of variables is very large, and one in which each of the many variables has a behavior which is individually erratic, or perhaps totally unknown [2]."

A system can be viewed as random because of numerous interactions of its constituent parts, where the properties of the system can *collectively* be understood by using probability and statistical mechanics. In other words, a system exhibits disorganized complexity if its constituent elements fluctuate at random.

Although systems exhibiting organized complexity can be confounding, their disorder differs from those characterized by disorganized complexity by virtue of their non-randomness:

> "…a very substantial number of variables is involved here, and they are all interrelated in a complicated, but nevertheless, not in helter-skelter fashion [2]."

The upshot is problems of organized complexity "…cannot be handled with the statistical techniques so effective in describing average behavior in problems of disorganized complexity [2]."

We propose that organized complexity can actually be thought of as a continuum rather than discrete values corresponding to random vs. non-random interactions. Furthermore, the magnitude of complexity for any entity or process is determined by the number and diversity of its constituent elements.

At one continuum extreme the constituent elements have zero diversity, *i.e.*, they are identical. Such a condition implies the configuration process has zero uncertainty, and therefore is entirely deterministic. Moreover, at this extreme—and only at this extreme—the magnitude of organized complexity is independent of the number of constituent elements, recalling the jigsaw puzzle consisting of 1000 identical pieces.

At the other continuum extreme an entity's elements have maximum diversity, *i.e.*, the configuration process has maximum uncertainty. In other words, configuration process outcomes are a uniformly distributed random variable. An entity in this condition is said to exhibit *disorganized* complexity, and the configuration process is equivalent to a fair coin toss if there are just two process outcomes.

In the continuum representation, disorganized complexity is a limiting case of organized complexity. Furthermore, the magnitude of organized complexity for a

given entity depends on both the number of its constituent elements and the uncertainty associated with configuring those elements.

The continuum as described above leads to a general definition of complexity:

> *The complexity of an entity or process describes the predictability in correctly configuring and/or ordering its constituent elements. Specifically, the magnitude of complexity is a function of the number and diversity of the constituent elements, and is inversely proportional to the probability of guessing their correct configuration in the absence of a priori knowledge of the relevant entity.*

Implicit in this definition is the critical notion that estimating complexity is an inherently probabilistic process. The blindfold in the jigsaw puzzle example was merely a prop to ensure the puzzle assembly process obeyed a uniform probability distribution. To be clear, the prop enabled us to quantify the probability of a correct puzzle configuration at the two extremes of the organized complexity continuum. However, the problem is less straightforward for intermediate continuum values, noting the configuration process is almost never entirely deterministic nor completely random.

Based on the definition, estimating the magnitude of complexity of an entity requires determining both the number of its constituent elements and their diversity. In theory it is trivial to determine the former by merely counting. Quantifying diversity is less obvious, and requires a more subtle approach.

## 3. Cybersecurity Complexity

According to the definition, the complexity of any entity is inversely proportional to the probability of guessing the correct configuration of its constituent elements. Therefore, if an IT environment is viewed as an entity consisting of a collection of cybersecurity risk factors, and each risk factor is either managed by a security control or not, the magnitude of *cybersecurity complexity* is inversely proportional to the probability of guessing the correct sequence of managed and unmanaged risk factors.

An information theoretic model of cybersecurity complexity has been proposed, and we closely follow the discussion in the cited reference to provide the background for the analyses that follow [1].

Assume IT vulnerabilities are either successfully managed or not, with probabilities p and 1-p, respectively. In other words, security risk management is assumed to be a binary stochastic process. A classic example of a binary stochastic process is a coin toss.

The uncertainty of any probability distribution is represented by its information entropy [3]. Information entropy is commonly referred to as Shannon

[1]Another application of entropy is in thermodynamics, where it is used to quantify uncertainty in the context of physical systems. However, it has been suggested that information entropy and thermodynamic entropy actually represent the same concept (see, "Information Theory and Statistical Mechanics" by E. T. Jaynes in "The Physical Review", 1957).

entropy or just "entropy," named after Claude Shannon, one of the originators of information theory.[1]

Entropy is defined as follows,

$$H = -\sum p_i \log_2 p_i \tag{1}$$

If the probabilities ($p$) are equal, (1) becomes,

$$H = -\log_2 p_i \tag{2}$$

The entropy of a binary stochastic process is shown in Figure 1.

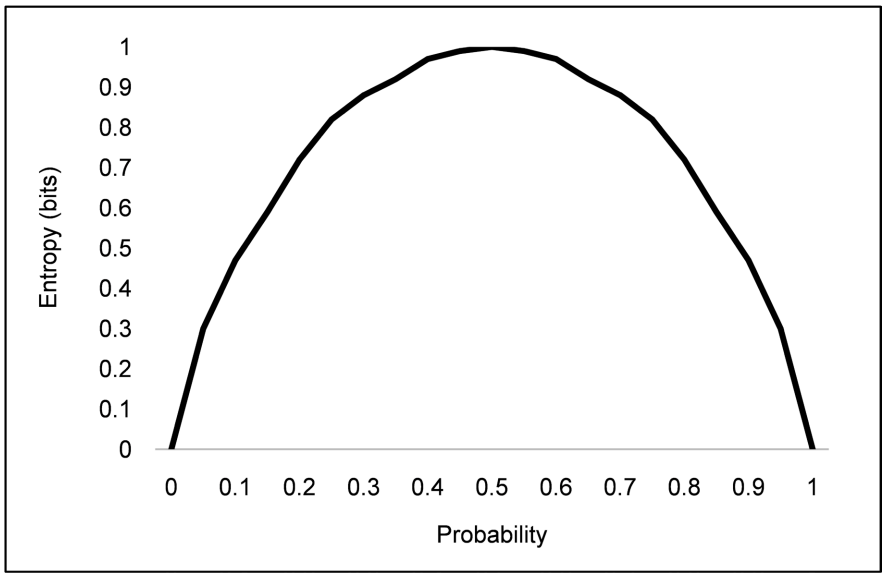


Figure 1. The entropy of a binary stochastic process.

The entropy of the probability distribution describing the toss of a fair coin is calculated as follows:

$$\begin{aligned} H &= -\left[ p(\text{heads}) \log_2 p(\text{heads}) + p(\text{tails}) \log_2 p(\text{tails}) \right] \\ &= -\left[ 0.5 \times (-1) + 0.5 \times (-1) \right] = 1 \text{ bit} \end{aligned}$$

In other words, it requires one "bit" of information to specify the outcomes of a coin toss when the two probabilities are equal and the logarithm is expressed in base 2. Moreover, the uncertainty of the coin toss process, *i.e.*, the entropy-per-coin toss, is a maximum when $H = 1$. In other words, the greater the process uncertainty the more information bits are required to specify the outcomes. Importantly, the greater the process uncertainty the more diversity of the outcomes resulting from multiple coin tosses.

According to the definition, the magnitude of complexity is inversely related to the probability of guessing the correct configuration of an entity's relevant constituent elements. This definition suggests that an estimate of complexity requires calculating the probability that a specific configuration of vulnerabilities results from security risk management. Since complexity in this context relates to cybersecurity, the relevant elements are the risk factors for information compromise,

*i.e.*, vulnerabilities. Therefore, it is necessary to estimate both the number of vulnerabilities and their diversity resulting from the application of security controls.

As noted above, estimating the number of vulnerabilities is an exercise in counting. Section 4 uses combinatorial methods to express the growth in the number of vulnerabilities as a function of network size, *i.e.*, the number of interconnected elements. General scaling relations are derived, but specifying the exact number of vulnerabilities in an actual IT environment is not a realistic objective.

Critically, estimating vulnerability diversity is possible using Figure 1 if and only if security risk management is assumed to be a binomially distributed random variable. This assumption is conceptually viable because complexity is defined probabilistically. Other probability distributions would be equally valid, but the significant conceptual leap is in assuming cybersecurity risk management is a stochastic process. A binomial probability distribution is a practical choice since it is simple and it obeys the Central Limit Theorem. If $p$ is not too close to either zero or one, *i.e.*, its limiting values, it quickly takes the form of a Gaussian distribution as the sample size increases.

Note that Figure 1 represents the entropy of *any* binomial probability distribution. In the case of a coin toss, $H$ specifies the entropy-per-toss. With respect to cybersecurity risk management, $H$ specifies the *entropy-per-vulnerability addressed.* In other words, if security risk management obeys a binomial stochastic process, each application of a security control to a vulnerability is mathematically equivalent to tossing a coin, where the heads and tails outcomes are replaced by managed and unmanaged vulnerabilities.

Therefore, the number of probable states resulting from the cybersecurity risk management process must equal $2^{VH}$, where $H$ is the entropy-per-vulnerability addressed and $V$ is the number of vulnerabilities ($V$) to which the security risk management process is applied. This quantity specifies the number of probable states resulting from cybersecurity risk management, which clearly increases exponentially with an increasing number of vulnerabilities.

In keeping with the qualitative definition of complexity, cybersecurity complexity $C$, is defined as follows,

$$C = 2^{-VH} \tag{3}$$

Specifically, $C$ represents the probability that an IT environment is in a particular state of the number of equally probable states resulting from cybersecurity risk management. The greater the number of these states the more complex the IT environment and the *smaller* the value of $C$, *i.e.*, the greater the cybersecurity unpredictability. When $H$ equals one, all *possible* states are equally probable, and such an environment exhibits maximum diversity. In any realistic IT environment, $V$ is likely to be a very large number, and H is always constrained to be between zero and one.

Unfortunately, it is impossible to precisely calibrate $H$ or to specify all the cybersecurity risk factors present in any realistic IT environment. However, it is possible to perform comparisons of one or both quantities for different scenarios.

Therefore, $C$ is most useful in characterizing *relative* complexity. To that end, we define the relative complexity $C_r$ to be the ratio of the number of probable states to the number of possible states with equal probability, *i.e.*, $H = 1$:

$$C_r = 2^{VH} / 2^V = 2^{V(H-1)} \tag{4}$$

$C_r$ expresses the *deviation* from a maximum complexity condition. Figure 2 illustrates the relative complexity for two IT environments containing $V = 10$ and $V = 5$ vulnerabilities.

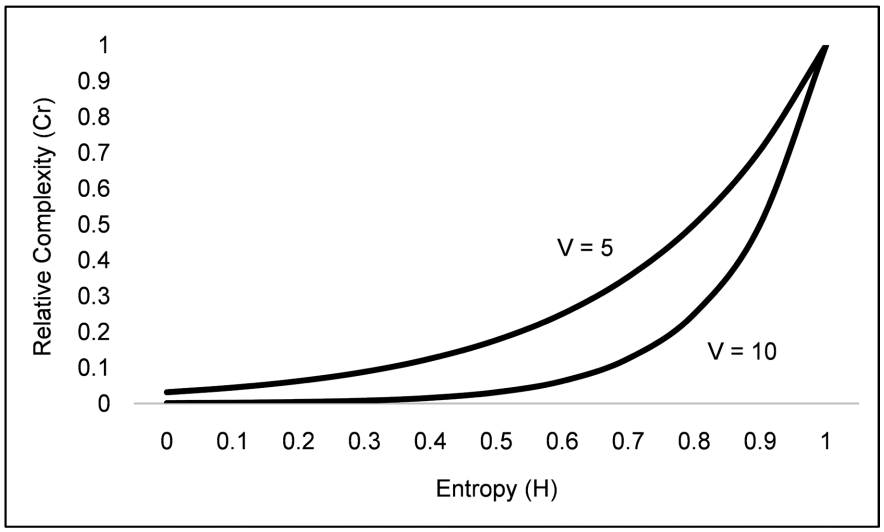


Figure 2. Relative complexity for IT environments with five and ten vulnerabilities.

Larger IT environments generally contain more vulnerabilities, but such environments would also be further from a maximum cybersecurity complexity condition for a given value of $H$ as shown in Figure 2. The disparity continues until $H$ equals one, wherein all IT environments are maximally complex, independent of their size.

A condition where $H$ equals one represents the theoretical limit for any IT environment. Such an environment would exhibit disorganized complexity, *i.e.*, maximum risk management uncertainty. This condition is arguably a worst-case scenario since the process of applying security controls to vulnerabilities would be equivalent to tossing a fair coin. Ideally security risk management introduces a bias such that uncertainty is actually minimized, *i.e.*, $H$ is close to zero.

Perhaps the most significant operational result of the previous discussion is that cybersecurity complexity increases *exponentially* with network size. This result is consistent with intuition since we would qualitatively expect the magnitude of complexity to somehow increase with size as noted in Section 2. This exponential dependence has strong operational/practical implications. Specifically, the model reveals the criticality of minimizing the total number of vulnerabilities in conjunction with maximizing IT environment uniformity on an enterprise scale. Therefore, security controls affecting IT environments at this scale such as those pertaining to security governance assume added significance.

However, and as was also noted in Section 2, size alone is not an accurate predictor of the magnitude of cybersecurity complexity. It turns out that network interconnectivity represents another risk-relevant feature of IT environments.

## 4. Network Interconnections

All networked elements can contain vulnerabilities as can the technologies that facilitate interactions between elements. In general, we would expect larger networks to contain more vulnerabilities, and therefore, the larger the network the greater the potential for a successful information compromise assuming all vulnerabilities are equivalent. The exponential dependence of complexity on $V$ means it is critical to determine how the number of vulnerabilities scales with network size in order to quantitatively evaluate cybersecurity complexity.

Note that increasing the number of network elements also increases the number of connections between elements, where the connection itself could contain multiple risk factors for information compromise due to the varied methods of communicating. Critically, these connections and the associated vulnerabilities increase *non-linearly* with the number of networked elements.

As a trivial example, consider an IT network consisting of four identical nodes. Since each node can connect to three of its neighbors (assuming self-connections are prohibited), there are twelve possible connections, noting half of the possible connections are redundant. Therefore, there are six unique combinations of connections arising from a four-node network. Figure 3 depicts such a network and the six possible connections.

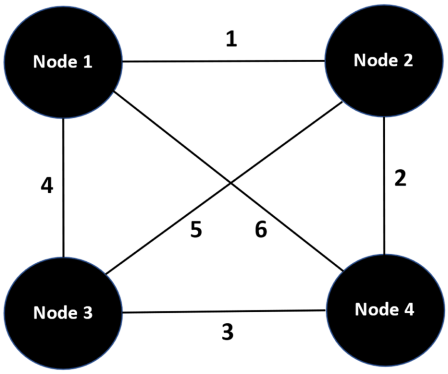


Figure 3. A four-node network and the possible interconnections.

Similarly, if there are five network nodes, every node can connect to each of the other four nodes. Therefore, the number of connections equals 10, *i.e.*, (4 connections/node × 5 nodes)/2. Figure 4 illustrates the number of possible interconnections as a function of the number of network nodes for up to 10 nodes.

We see that the number of possible interconnections scales quadratically, *i.e.*, an exponent of 2, with the number of nodes. In this case we have assumed all nodes are identical, which is not necessarily true in an actual IT environment. In

view of the exponential dependence of $C$ on the number of vulnerabilities per (5), this non-linear relationship has significant implications to cybersecurity complexity, and hence to the potential for a successful information compromise. Therefore, in order to fully characterize cybersecurity complexity, it is necessary to determine the precise relationship between network size and the number of vulnerabilities resulting from network interconnections.

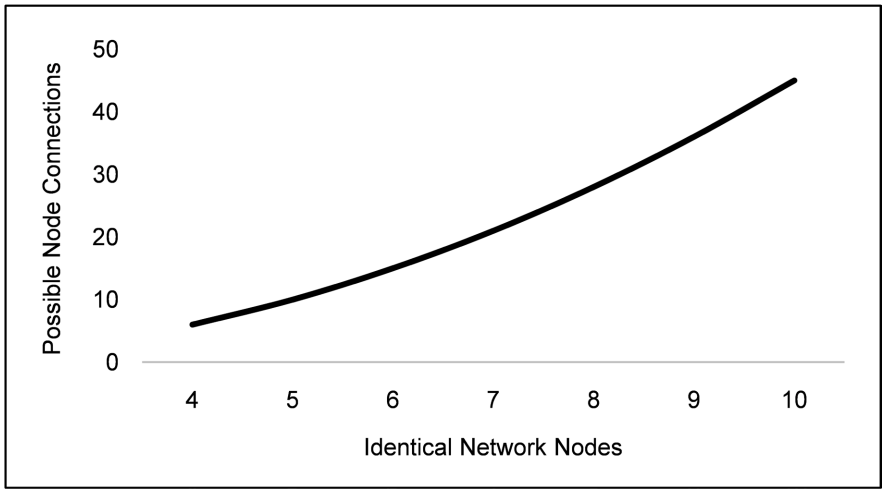


**Figure 4.** The effect of adding network nodes on the possible number of node connections.

## 5. IT Vulnerability Scaling Relations

The scale used to assess cybersecurity risk will significantly affect the result. As a trivial example, examining single endpoints for vulnerabilities will yield a different perspective than an assessment of the entire IT environment.

Furthermore, the potential for a successful information compromise is surely scale-dependent. The larger the IT environment the more targets for attack, opportunities for data leakage, IT users to be phished, connections made to malicious web sites, vulnerabilities within endpoints, mismanaged security controls, etc. Cybersecurity professionals refer to the "attack surface" in describing the breadth of exposed IT environment elements. In general, the larger the attack surface the greater the potential for a successful information compromise.

As discussed in Section 4, the number of possible interactions between network elements increases with network size. We note that the network includes all devices joined to the network, and there are numerous ways of adding and/or managing such devices. These include Mobile Device Management, Network Access Control, VPN, Cloud Directory Services, Domain Joining (e.g., Active Directory), Endpoint Security Solutions, and custom scripts and automation tools.

Once devices are joined to the network, interactions between devices can be achieved via multiple methods, e.g., email, SMS, communication applications, and each of these methods potentially contains vulnerabilities. Recall (3) revealed the exponential dependence of cybersecurity complexity on the total number of vulnerabilities. Therefore, specifying the scale-dependence, *i.e.*, a scaling relation, of

cybersecurity complexity on the number of possible vulnerabilities would directly relate to the potential for a successful data breach.

Identifying such a relation requires a more nuanced view of IT vulnerabilities, which we claim are broadly divided into two groups: "local" and "non-local." We define local vulnerabilities to be exploitable features of a device, system, or application that are confined to the relevant device, system, or application. These vulnerabilities might be managed by applying a patch that is replicated across all instances within an IT environment.

Examples of local vulnerabilities are those rated by the Common Vulnerability Scoring System [4]. This system uses six metrics to compute exploitability and impact sub-scores, which are concatenated to produce the CVSS "vector" for a given vulnerability. Note this locality condition does not imply that local vulnerabilities cannot be used to compromise other IT environment elements or can't be exploited in combination with other local vulnerabilities.

In contrast, non-local vulnerabilities result from the *interactions* of elements, *i.e.*, nodes, within an IT environment. For example, a non-local vulnerability might arise as a result of a technology configuration or setting of an endpoint running an application that facilitates a user-managed process or workflow. The latter might send the resulting data to another endpoint operated by a different IT user. Such actions are repeated numerous times across the environment.

We next derive scaling relations relating the number of local and non-local vulnerabilities to the size of an IT environment, such as the generic environment depicted in Figure 5.

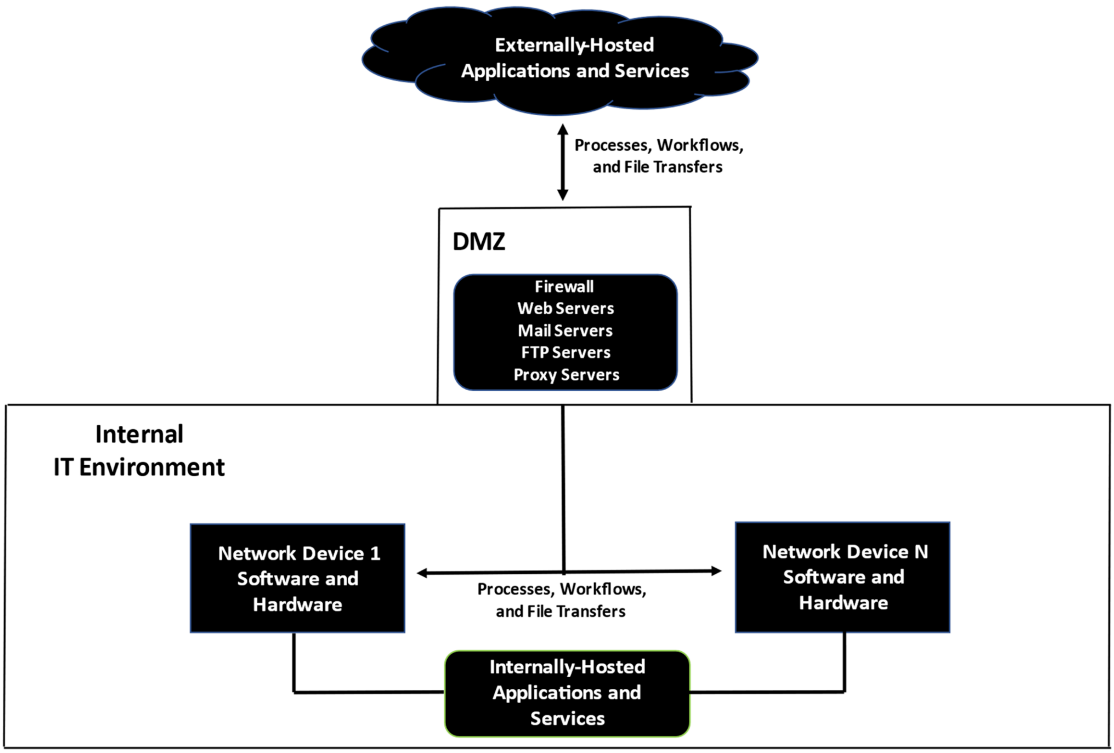


Figure 5. Generic IT environment.

The number of internal network devices, *i.e.*, nodes, can vary significantly, and that variability is reflected in the indices labeled 1 to some arbitrary number $N$. If there are $N$ nodes each having $M$ local vulnerabilities, the number of local vulnerabilities, $V_{\text{local}}$, equals $M \times N$. In this case there is a *linear* relationship between $T$ and $M$ with $N$ fixed, or between $V_{\text{local}}$ and $N$ with $M$ fixed. $V_{\text{local}} = M \times N$ is the scaling relation for local vulnerabilities in an arbitrary IT environment.

Therefore, applying security updates, patches, anti-virus, and similar techniques to $N$ nodes remediates $M \times N$ vulnerabilities. The linearity of $V_{\text{local}}$ dictates the requirement for remediation as a function of network size since size relates directly to $N$. Note the relationship becomes more complicated if either $M$ or $N$ is permitted to change.

Non-local vulnerabilities are associated with multiple nodes and depend on the interactions between nodes. Here security controls manage cybersecurity risk at the network level rather than the node level. We note that the actions of IT users, *i.e.*, humans, are frequently involved in node interactions. Therefore, non-local vulnerabilities can arise from user behavior. The key point is that non-local vulnerabilities are due to node interactions rather than the node itself.

If there are $N$ nodes, where a node has $M$ methods to exchange data, each method can be considered a non-local vulnerability since it specifically relates to node interactions rather than the node technology. Therefore, the number of non-local vulnerabilities depends on the number of interacting nodes.

For example, if there are $N$ interacting nodes, the total number of non-local vulnerabilities $V_{\text{non-local}}$, equals $MN(MN\text{-}1)$.[2] Critically, this expression represents a *non-linear* relationship, where for large $M$ or $N$, $V_{\text{non-local}} \sim (MN)^2$. Note this formulation assumes that any of the $M$ communication methods can communicate with any other method. It also treats each direction of communication as a unique vulnerability. Otherwise, $V_{\text{non-local}} = MN(MN\text{-}1)/2$.

This logic could be extended to an arbitrary number of interacting nodes. If the number of nodes required for a particular interaction is $Q$, the following relation holds:

$$V_{\text{non-local}} = MN\left(MN-1\right)\left(MN-2\right)\left(MN-3\right)\cdots\left(MN-\left(Q-1\right)\right),$$

which equals $(MN)^Q$ if $Q$ is $\ll MN$, and is essentially a first-order Stirling approximation.[3]

As an example of a scenario where non-local vulnerabilities arise, consider two IT users working on a shared file via their respective machines, *i.e.*, nodes, using a Cloud-hosted application. Each user must contribute work following successful authentication. Email communications are exchanged among the users, and progress is logged locally and in the Cloud. Verbal communications among IT users might also occur. Note that although two users might initially be engaged in the process, the number of users and accompanying technologies might change as

[2]Note this formulation includes the mathematically possible but physically unrealistic cases where nodes can communicate with themselves.

[3]Stirling's approximation is given by $n! \sim (2\pi n)^{1/2}(n/e)^n$.

work progresses.

A more general formula for the number of vulnerabilities per interaction, where self-interactions are allowed, is given by (5). Here $i$ represents the number of nodes required for a given information exchange, and the exclamation point is the factorial sign, where $N! = N \times (N-1) \times (N-2) \times (N-3) \times \cdots \times (N-(N-1))$.

$$V(i, M(i), N(i)) = \frac{(M(i)N(i))!}{(M(i)N(i)-i)!} \tag{5}$$

The total number of vulnerabilities would be determined by summing over the variable $i$, up to whatever maximum value is applicable to the particular IT environment. Note that $M$ and $N$ have become functions of $i$, which indicates the number of methods of exchanging data as well as the number of nodes utilizing such methods can change based on the number of simultaneous interactions.

Local vulnerabilities as represented by (5) in the special case where $i = 1$, *i.e.*, a node interacting with itself, plus non-local vulnerabilities, represent the more general case of $i > 1$. The number of methods to exchange data ($M$) is almost certainly dependent on i in any realistic IT network. However, the number of nodes ($N$) is more likely to be independent of $i$ over risk-relevant time scales since nodes can potentially be communicating in pairs, trios, etc.

Each direction of node interaction is treated as an independent vulnerability. For example, node A sending information to node B represents a distinct vulnerability relative to node B sending information to node A. Therefore, the above formula only reflects permutations and not combinations. If these connections were treated as a single vulnerability, (1) would be divided by $i$ factorial.

Finally, (6) is a general expression for the total number of vulnerabilities per interaction, where self-interactions are not allowed:

$$V(i, M(i), N(i)) = \frac{M^i(i)(N(i))!}{((N(i)-i))!} \tag{6}$$

Again, the total number of vulnerabilities in a given environment would be determined by summing (6) over $i$, up to the maximum value that applies to the IT environment under consideration.

Note the case where $i = 2$, *i.e.*, two communicating nodes, would be a common scenario, where $M(i)$ decreases significantly for higher values of $i$. $(MN)^2$ would be the typical result irrespective of whether self-interactions were allowed or not, and this quadratic expression is clearly non-linear.

The implication of (5) and (6) is that non-local vulnerabilities scale *non-linearly* with IT environment size. This non-linearity suggests that non-local vulnerabilities exert an exponentially increasing effect on cybersecurity complexity as the environment size increases.

## 6. Reducing Cybersecurity Complexity

The previous discussion naturally leads to a discussion of the operational implications of the results. Specifically, the non-linear dependence of vulnerabilities on

network size clearly means decreasing the number of non-local vulnerabilities is key to reducing cybersecurity complexity. Decreasing the number of non-local vulnerabilities entails reducing the number of network interactions. In this section we derive vulnerability scaling relations for specific topologies that reduce the number of possible network interactions, which exponentially decreases cybersecurity complexity in accordance with (3).

The simplest topology is the use of a central network node that mediates interactions between non-central or "standard" nodes. Note that all such interactions require three nodes in total: two standard nodes and a central node. The central node is sometimes referred to as a proxy, where proxies are commonly used to facilitate secure intra-network communications by performing packet inspection on data exchanges. Figure 6 illustrates the intra-network interactions for this network topology, noting that protecting the central node is paramount given its role in facilitating network interactions.

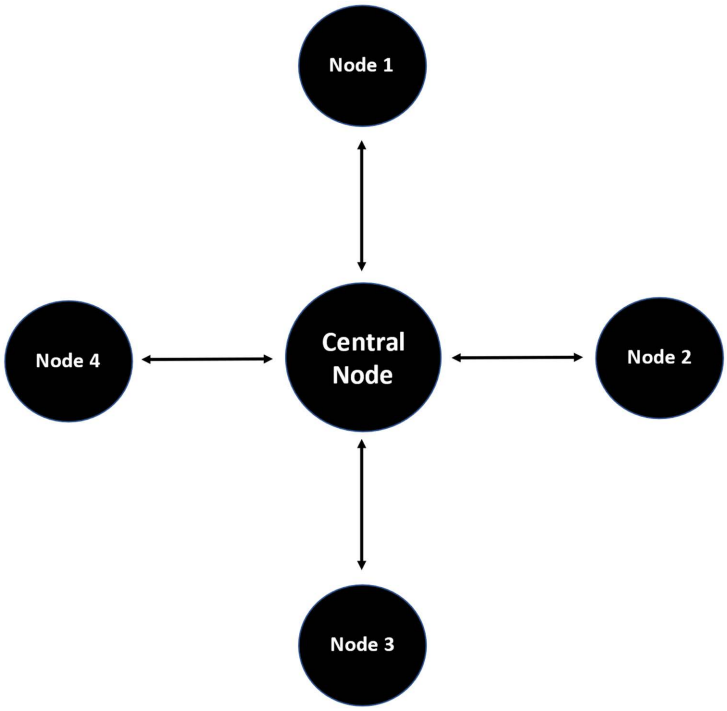


Figure 6. Central node network interactions.

Assume a standard node can interact with a central node via $M$ methods, and there are $S$ secure methods for the central node to interact with a standard node, where $S \leq M$. Recall $N$ is the number of nodes. All interactions occur between node pairs, *i.e.*, $i = 2$. Therefore, the scaling relation for the number of vulnerabilities ( $V$) is as follows:

$$V(i=2) = M(N-1) + S(N-1) = (M+S)(N-1) \tag{7}$$

We see that the number of non-local vulnerabilities scales linearly with $N$, *i.e.*, $V(2) \approx (M+S)N$. If $M = S$, then $V(2) \approx 2MN$, and in that case $V(2)$ scales linearly

with both $N$ and $M$.

A second topology to attenuate the non-linear growth in vulnerabilities consists of standard nodes separated into $D$ divisions. Each division has a "special" node that mediates any interaction across divisions. All standard nodes can interact with any other standard nodes within the same division. Figure 7 illustrates the intra and inter-divisional interactions for this topology.

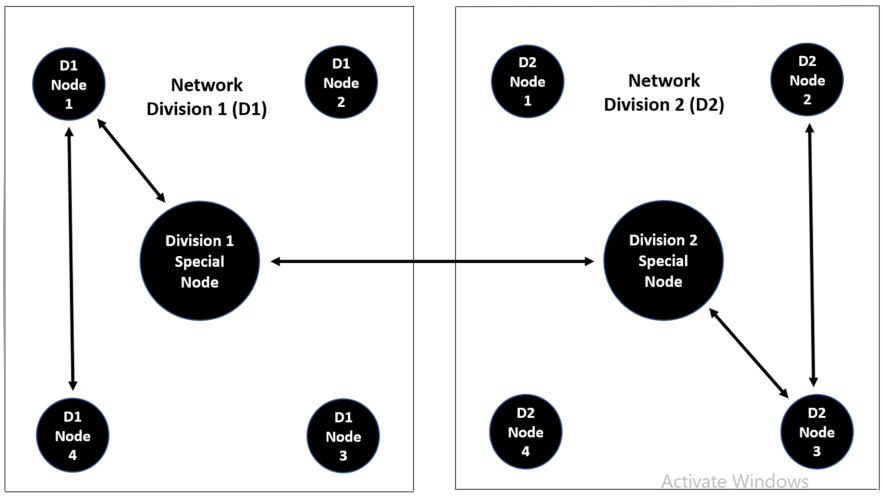


Figure 7. Intra and inter-divisional network interactions.

Assume there are $M$ methods for a standard node to interact with nodes (both standard and special) within the same division. Furthermore, there are $S$ secure methods for special nodes to communicate with special nodes in other divisions, where $S \le M$. This configuration reduces all interactions to pairs of nodes ($i = 2$) only when interacting across divisions so that the number of vulnerabilities becomes,

$$V(i) = DM^i \frac{\left(\frac{N}{D}\right)!}{\left(\frac{N}{D}-1\right)!} + S^i \frac{(D)!}{(D-1)!} \tag{8}$$

However, most interactions require pairs of nodes, *i.e.*, $i = 2$, and we therefore focus on the special case, $V(2)$:

$$V(2) = M^2 N\left(\frac{N}{D}-1\right) + S^2 D(D-1) \tag{9}$$

Taking the derivative of (9) with respect to $D$ to minimize $V(2)$ and solving for $D$ results in a cubic equation that expresses the number of divisions that minimizes the number of non-local vulnerabilities:

$$D^3 - \frac{D^2}{2} - \frac{M^2 N^2}{2S^2} = 0 \tag{10}$$

$N$ is generally large for realistic scenarios, so $D \gg 1$. Therefore, we can ignore the $D^2$ term to arrive at an approximate solution:

$$D^* \approx \left( \frac{M^2 N^2}{2S^2} \right)^{\frac{1}{3}} \tag{11}$$

Substituting (11) into (9) to obtain the minimized number of non-local vulnerabilities yields the following expression:

$$V(2)_{\min} = \frac{3}{2^{\frac{2}{3}}} M^{\frac{4}{3}} S^{\frac{2}{3}} N^{\frac{4}{3}} - M^2 N - S^2 \left( \frac{M^2 N^2}{2S^2} \right)^{\frac{1}{3}} \tag{12}$$

If $M = S$, then (12) becomes,

$$V(2)_{\min} = M^2 \left( \frac{3}{2^{\frac{2}{3}}} N^{\frac{4}{3}} - N - \frac{N^{\frac{2}{3}}}{2^{\frac{1}{3}}} \right) = M^2 N^{\frac{4}{3}} \left( \frac{3}{2^{\frac{2}{3}}} - \frac{1}{N^{\frac{1}{3}}} - \frac{1}{2^{\frac{1}{3}} N^{\frac{2}{3}}} \right) \tag{13}$$

From (12) and (13) we see that non-local vulnerabilities scale above linearly but still sub-quadratically with $N$, *i.e.*, $V(2) \sim N^{4/3}$, and scale quadratically with $M$, if $M = S$.

Applying a Taylor series expansion of (5) around the value $D^*$ to lowest order, where $D = D^* + x$, and $x \ll D^*$, we see that $V(2)$ displays a strong quadratic dependence on both $x$ and $S$:

$$V(2) \approx V(2)_{\min} - S^2 x + 3S^2 x^2 \tag{14}$$

This result shows a significant increase in the number of vulnerabilities if the number of divisions ($D$) is significantly different than $D^*$. In addition, it suggests that minimizing $S$, the number of interactions between special nodes, helps reduce the impact of such a deviation.

Consider the case where $M = S = 1$, and $N = 100$. The value of $D^*$ is roughly equal to 17, with a minimized number of vulnerabilities equal to approximately 760. The two curves in **Figure 8** are plots of (14) and (9) as a function of $x$. The difference in the two curves for small deviations in $x$ reveals the impact on the number of vulnerabilities when straying from $D^*$, which in turn highlights the importance of keeping the number of central nodes at or near the minimum number of divisions. Both curves demonstrate the sharply quadratic dependence on the number of vulnerabilities as a function of $x$.

Note the dependence on $N$ in (14) exists within $V(2)_{\min}$, per (13). Expanding to orders above quadratic (*i.e.*, cubic terms and above) would also introduce a dependence on $N$ in the correction terms. The significant overlap of the two curves in **Figure 8** adequately captures the response of the network to changes in $D$, thereby obviating the need for higher order analysis.

To illustrate the potentially dramatic effect of reducing the number of network nodes on the number of possible vulnerabilities, we first set $M = 1$, $i = 2$, $N = 100$, and compare the results with $N = 200$ for various scaling relations. **Table 1** shows the impact of linear versus sub-quadratic and quadratic scaling relations.

The effect of doubling $N$ is to double the number of non-local vulnerabilities if a linear scaling relationship exists. In contrast, doing the same for sub-quadratic

and quadratic scaling relations increases the number of non-local vulnerabilities by a factor of 2.5 and 4 respectively. In other words, a quadratic scaling relationship has a disproportionately greater impact on cybersecurity complexity than either a sub-quadratic, e.g., $N^{4/3}$, or linear scaling relationship. This finding argues strongly for implementing network topologies exhibiting sub-quadratic or (even better) linear vulnerability scaling, all other considerations being equal.

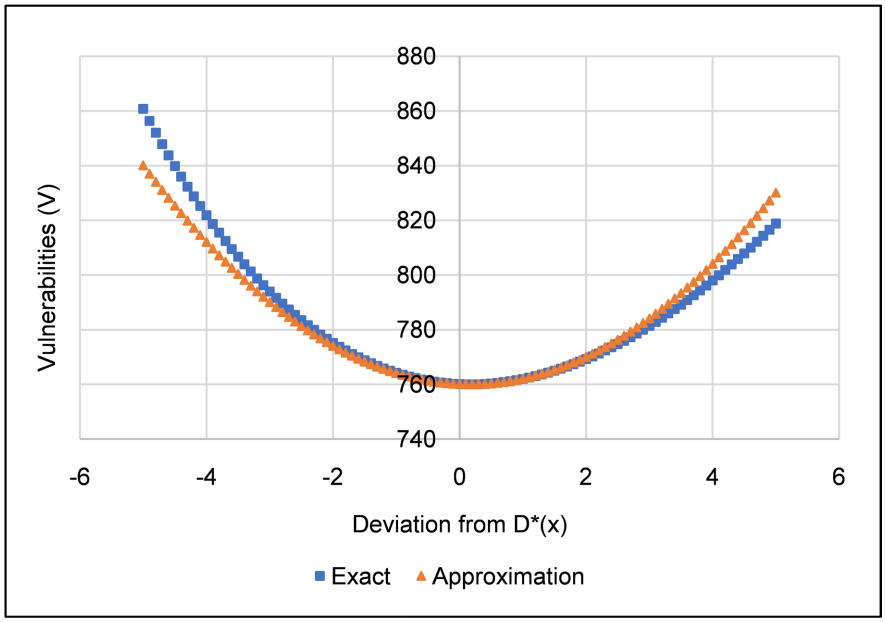


**Figure 8.** The effect on the number of vulnerabilities as a result of deviating from a minimized number of divisions/central nodes.

**Table 1.** The effect of doubling the number of nodes on the number of non-local vulnerabilities.

| Scaling Relation | No. of Non-Local Vulnerabilities If $N = 100$ | No. of Non-Local Vulnerabilities If $N = 200$ |
|---|---|---|
| Quadratic | 10,000 | 40,000 |
| Sub-Quadratic | 880 | 2200 |
| Linear | 200 | 400 |

Note that special nodes represent concentrations of risk since they represent the sole communication agents across divisions. Although special nodes are not required to mediate intra-divisional interactions such as the central nodes of **Figure 6**, their compromise could eventually result in the compromise of all standard nodes within a division following a single inter-divisional interaction.

It is interesting to note that the standard assumption is that increasing network segmentation actually increases complexity. However, its precise effect is to increase *administrative* complexity. In that context, complexity relates specifically to the additional overhead required to sustain multiple network segments, e.g.,

VLANs, subnets.

In fact, the benefit of adding divisions beyond $D^*$ exponentially diminishes due to the potential vulnerabilities introduced by adding central nodes. Therefore, $D^*$ might be considered a demarcation between administrative and cybersecurity complexity. In other words, below $D^*$, the principal contributor to IT environment complexity is administrative complexity. Above $D^*$, cybersecurity complexity dominates, where the dominance increases exponentially with increasing $D$ since administrative complexity presumably scales linearly with the number of possible interactions.

## 7. Categorizing Security Controls

The previous analysis coupled with the model for cybersecurity complexity in Section 3 helps explain why certain security controls are more or less effective. Recall that cybersecurity complexity is a function of both the number of vulnerabilities ($V$) and the entropy per vulnerability ($H$). These parameters facilitate the organization of cybersecurity complexity controls according to whether they reduce $V$, $H$, or both.

The first category is a security control that reduces the number of local vulnerabilities. This type of control is applied to the systems, devices, and applications where such vulnerabilities reside. They are generally less effective at addressing macroscopic security phenomena since such phenomena are most affected by interactions between network elements. As discussed in Section 6, a network topology that restricts communication between nodes exponentially reduces the number of non-local vulnerabilities by decreasing the number of possible interactions, with attendant implications to cybersecurity complexity.

The second category of security control reduces the entropy-per-vulnerability ($H$). This category limits other types of interactions within the network, where the net effect is to drive an IT environment away from a condition of *disorganized* complexity. An example is an information security policy, which reduces $H$ by fiat, and is intended to limit interactions affecting data exchanges on an enterprise scale. Importantly, IT users frequently initiate such interactions, which only emphasizes the importance of enforcing information security policy requirements since human behavior, exacerbated by cybersecurity complexity, is frequently the root cause of cybersecurity incidents.

Another example of the second category of security controls is a policy of Zero-Trust. Here interactions between nodes are limited by restricting network communications according to pre-approved parameters. Therefore, this policy effectively reduces the entropy that naturally accompanies such interactions, which has a salutary effect on cybersecurity complexity.

A third control category is one that limits both $V$ and $H$. For example, restricting inter-node communication via the network topology decreases $H$ by reducing security risk management uncertainty in addition to decreasing the number of possible non-local vulnerabilities as a function of network size. Since $C$ is an

exponential function of both $V$ and $H$ according to (5), the potential effect of implementing the topologies depicted in Figure 6 and Figure 7 is dramatic.

Practical applications of these results follow immediately. First, it provides a quantitative justification for reducing $H$ via security governance, which is a security control that by definition affects organizations on an enterprise scale. In that vein, maximally reducing $H$ by enhancing security governance suggests the need to develop information security policies that address the largest number of potential use cases. One example might be a policy structure organized according to the phases of the information lifecycle, e.g., information creation, transmission, storage, and deletion, and the sources and types of information relevant to the organization in question.

A second practical application of these results is the increased use of proxies as communication intermediaries between groups of networked nodes. As noted above, the results argue strongly for implementing network topologies exhibiting sub-quadratic, or ideally, linear vulnerability scaling, all other considerations being equal. Implicit in these results is the very practical recommendation that nodes functioning as intermediaries will require stringent security monitoring because of their central role in intra-network communication.

A third practical application is the control categorization itself. Part of the complexity inherent to cybersecurity is the lack of a formal cybersecurity risk management taxonomy. We allege without proof that any cybersecurity risk management strategy that facilitates a more structured approach to applying security controls to risk factors possesses incremental advantages relative to strategies that do not, all other factors being equal.

Finally, one could argue this model of complexity provides a conceptual basis for implementing a policy of Zero-Trust by assigning a value of uncertainty to the process of applying security controls to risk factors. Although not a practical result *per se*, the justification for implementing Zero-Trust on an enterprise scale immediately becomes rigorous and compelling as a result.

## 8. Conclusions

The effect of complexity on cybersecurity risk management has historically been difficult to describe let alone quantify. Pinpointing its specific contribution to cybersecurity risk will likely never be possible, in part because its effects are only manifest when a myriad of interacting network elements and inter-related phenomena converge in reducing the effectiveness of cybersecurity risk management. The difficulty inherent in quantifying the effectiveness of security risk management ultimately stems from the inability to perform statistical inference.

In its absence, it is sometimes possible to leverage concepts from other disciplines, and thereby invoke useful analogies. Statistical physics is one such discipline, where the uncertainty that arises from numerous interacting network elements is modestly reminiscent of the uncertainty resulting from large collections of non-interacting atoms. Importantly, the uncertainty common to each scenario

suggests a stochastic view of cybersecurity risk management has applicability. Furthermore, defining complexity probabilistically—a rational if expedient approximation—naturally leads to a probability distribution that describes security risk management outcomes.

As one might expect, this probabilistic definition was crucial to justifying the stochastic analyses that followed. Moreover, alternative definitions of complexity would almost certainly yield different results. That said, probability theory is what facilitated generalizations about complexity across disparate IT environments. Cybersecurity complexity has heretofore defied characterization precisely because exclusively deterministic models of cybersecurity risk management preclude such generalizations. Therefore, it is unclear whether non-probabilistic definitions of complexity would actually produce fruitful results.

However, pushing the analogy between cybersecurity and statistical physics too far would be a mistake since IT environment elements do not undergo random fluctuations. The bottom line is a stochastic formulation of security risk management is imperfect, but it is likely the only way to quantify complexity in this context. Furthermore, it is consistent with the qualitative definition of complexity.

In addition, we admittedly adopted a coarse-grained approach, wherein all network elements were assumed to be functionally equivalent. This simplification was key to showing the strong dependence of complexity on interactions between nodes. It was also relevant to demonstrating the non-linearity of non-local vulnerabilities as a function of network size, which in turn revealed the salutary effect of reducing intra-network interactions.

Two such strategies emerged from this analysis, where each strategy was based on network topologies that constrained the possible interactions between nodes. One topology yielded a linear scaling of vulnerabilities between pairs of nodes. The other resulted in non-linear but still sub-quadratic scaling of vulnerabilities between pairs of nodes. Each strategy has both operational and security implications requiring consideration prior to implementation.

These strategies won't eliminate cybersecurity complexity since local vulnerabilities will be largely unaffected, and non-local vulnerabilities will never entirely disappear. Furthermore, accurately predicting the effect of any risk management strategy without the benefit of statistical inference is not a realistic objective. However, perhaps as important as identifying a specific remedy is to address macroscopic security phenomena as part of the organizational strategy, and thus expand the focus of remediation efforts beyond local vulnerabilities.

We note the approach to characterizing cybersecurity complexity as described herein differs significantly from existing methodologies. For example, a method such as FAIR represents a high-level framework for assessing security risk rather than a model of a principal driver of cybersecurity risk that facilitates quantitative risk analyses.

In a similar vein, although frameworks such as NIST and CIS are helpful in assessing the effectiveness of cybersecurity risk management, they too do not

model cybersecurity risk nor do they attempt to quantify the effect of a risk-relevant organizational feature such as complexity. Traditional cybersecurity risk assessments focus on individual vulnerabilities that rely on commercial scanning applications, e.g., Tenable, Qualys, and Rapid 7, which are necessary but not sufficient to assess cybersecurity risk on an enterprise scale.

Future work suggested by these results might involve exploring complexity beyond network nodes, such as examining its effect on identity and access management (IAM) boundaries. Although our analysis has inherent implications for a policy of Zero-Trust, a more formal analysis of complexity in that specific context might also be worthwhile. Finally, the connection between information entropy and so-called promise theory might inspire additional insights into cybersecurity complexity, and therefore could be worthy of investigation [5].

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Young, C. (2022) Cybercomplexity, a Macroscopic View of Cybersecurity Risk. Springer Nature, 59-106.

[2] Weaver, W. (1948) Science and Complexity. *American Scientist*, **36**, 536-544.

[3] Shannon, C. and Weaver, W. (1949) The Mathematical Theory of Communication. The University of Illinois Press.

[4] National Vulnerability Database (NIST). https://nvd.nist.gov/vuln-metrics/cvss

[5] Wikipedia, Promise Theory. https://en.wikipedia.org/wiki/Promise_theory