Defense Strategies for Cloud Computing Multi-Site Server Infrastructures

Nageswara S.V. Rao
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA
raons@ornl.gov

Chris Y. T. Ma
Hang Seng Management College
Hong Kong
chris.ytma@gmail.com

Fei He
Texas A&M University
Kingsville, TX, USA
fei.hei@tamuk.edu

ABSTRACT

We consider cloud computing server infrastructures for big data applications, which consist of multiple server sites connected over a wide-area network. The sites house a number of servers, network elements and local-area connections, and the wide-area network plays a critical, asymmetric role of providing vital connectivity between them. We model this infrastructure as a system of systems, wherein the sites and wide-area network are represented by their cyber and physical components. These components can be disabled by cyber and physical attacks, and also can be protected against them using component reinforcements. The effects of attacks propagate within the systems, and also beyond them via the wide-area network. We characterize these effects using correlations at two levels using: (a) aggregate failure correlation function that specifies the infrastructure failure probability given the failure of an individual site or network, and (b) first-order differential conditions on system survival probabilities that characterize the component-level correlations within individual systems. We formulate a game between an attacker and a provider using utility functions composed of survival probability and cost terms. At Nash Equilibrium, we derive expressions for the expected capacity of the infrastructure given by the number of operational servers connected to the network for sum-form, product-form and composite utility functions.

CCS CONCEPTS

•Computing methodologies →Modeling; •Networks →Cyber-physical networks;

Keywords and phrases: cloud server infrastructure, networked systems, composite utilities, aggregated correlation function, game theory, Nash Equilibrium

1 INTRODUCTION

Big data applications over cloud computing infrastructures may span across multiple server sites, which are connected over wide-area networks. In these infrastructures, the wide-area network plays a critical, asymmetric role: its failures render the servers unreachable even if they are operational, and in extreme cases can render the entire infrastructure unavailable to users. We represent such an infrastructure by a system of systems consisting of the sites, S_i, $i = 1, 2, \ldots, N$ and the wide-area network S_{N+1} [15]. The sites are complex systems, each consisting of several discrete cyber components, including servers and network devices, and physical components, including site network fiber connections and Heating, Ventilation and Air Conditioning (HVAC) system.

A key performance metric for this infrastructure is the *capacity* given by the number of servers that are operational and accessible over the wide-area network. A variety of cyber and physical attacks can be launched on its components that degrade the capacity in different ways. The servers are accessible to users over the network, which makes them vulnerable to cyber attacks that can divert their processing power or simply crash them. In contrast, network routers are geographically dispersed with a restricted access by network administrators. Thus, cyber attacks on them require different techniques compared to server attacks, and have different effects on the capacity. Successful attacks on routers can disconnect significant portions of the network, rendering the servers at disconnected sites unavailable to users. Attacks on network elements at the sites such as LAN switches and border routers have similar but somewhat localized degradation effects. The increasing deployment of network control apps for site HVAC systems, particularly on smart phones, makes them vulnerable to cyber attacks, which for example can increase the facility temperature to trigger server shutdowns. Physical attacks in the form of fiber cuts and cooling tower degradations represent different attack vectors that degrade this infrastructure; however, they require a proximity access by an attacker. For example, fibers connecting server sites to gateway routers and in between wide-area routers may be physically cut, thereby making sites and portions of the network inaccessible to users. Degradations of HVAC cooling towers, which are typically in open areas outside the sites, can lead to the shutdown of all site servers and network devices.

Various component reinforcements may be put in place to protect against the above attacks, including replicating servers and routers for fail-over operations, and installing redundant fiber lines to the sites and between wide-area network router locations. While such reinforced components can survive direct attacks, the servers may still be unavailable to users due to propagative effects of attacks on other components. For instance, even if all servers at a site are hardened against cyber attacks, they can all be made unavailable, for example, by cutting the fiber connections to the site with a single physical attack, or by bringing down the HVAC system by a single cyber attack. Non-reinforced components, on the other hand, will be disabled by direct attacks. The reinforcements and attacks incur costs to the provider and attacker, respectively, and their corresponding benefits depend not only on the components

ICDCN '18, January 4-7, 2018, Varanasi, India

DOI: 10.1145/nnnnnnn.nnnnnnn

but also on various correlations between components and systems, due to the propagation of disruptions within the sites and between them over the network.

Let n_i denote the number of components of S_i of which y_i and x_i denote the number of components attacked and reinforced, respectively. Let P_i be the survival probability of S_i, and P_I be the survival probability of entire infrastructure. The *expected capacity* of the infrastructure is the expected number of available components, given by

$$N_I = \sum_{i=1}^{N} n_i P_i,$$

which reflects the part of infrastructure that survives the attacks. Also, let S_{-i} denote the infrastructure without S_i, and P_{-i} be its survival probability. The relative importance of S_i is captured by the *aggregate failure correlation function* C_i [16] given by the failure probability of S_{-i} given the failure of S_i. The asymmetric role of the network is specified by two conditions [15]: (a) $C_{N+1} = 1$ indicates that network failure will disrupt the entire infrastructure, and (b) $C_i = 0$, for $i = 1, 2, \ldots, N$, indicates that disruptions of individual systems are uncorrelated. The correlations between components of individual systems are captured by simple first-order differential conditions on P_i [16]. This two-level characterization helps to conceptualize the basic correlations in this infrastructure, and provides insights into the needed defense strategies by naturally "separating" the system-level and component-level aspects.

A game between an attacker and a provider involves balancing the costs of attacks and reinforcements of systems, given by $L_A(y_1, \ldots, y_{N+1})$ and $L_D(x_1, \ldots, x_{N+1})$, respectively, with the survival probability of the infrastructure P_I. The *sum-form utility function* is given by

$$\begin{aligned} U_{D+} = & -[P_I(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1})]\, g_D \\ & + L_D(x_1, \ldots, x_{N+1}), \end{aligned}$$

which will be minimized by the provider, and the scalar $g_D \geq 0$ represents the benefit of keeping the infrastructure operational. The Nash Equilibrium (NE) is determined by the optimization of the utility functions by the defender and attacker, which in turn determines the capacity of the infrastructure. At NE, We derive the expected capacity for sum-form utility function, which indicates that higher gain g_D leads to lower number of operational and accessible servers. It also provides additional insights, for example, faster than linear C_i leads to lower number of available servers. We carry out similar analysis using a product-form utility function that represents a different cost-benefit trade-off compared to sum-form utility function (Section 4.2). Additionally, we also consider composite utility functions that subsume both sum-form and product-form utilities as special cases (Section 4.2).

The organization of this paper is as follows. We briefly describe the related work in Section 2. In Section 3, we describe the multi-site cloud computing infrastructure model of [15] expanded to include HVAC components, along with the aggregate correlation function and differential conditions on system survival probabilities. We present our game-theoretic formulation in Section 4 using sum-form and product-form utilities in Section 4.1, and using composite utility functions in Section 4.2, wherein we derive NE conditions and estimates for the system survival probabilities. Estimates for expected capacity are discussed in Section 5. We present conclusions in Section 6.

2 RELATED WORK

Critical infrastructures of power grids, cloud computing, and transportation systems rely on communications networks for connecting their constituent systems. These infrastructures are under increasing cyber and physical attacks, which the providers must counter by applying defense measures and strategies. Game-theoretic methods have been extensively applied to develop the needed defense strategies [1, 2, 10]. A comprehensive review of the defense and attack models in various game-theoretic formulations has been presented in [9]. Recent interest in cyber and cyber-physical systems led to the application of game theory to a variety of cyber security scenarios [10, 20], and, in particular, for securing cyber-physical networks [3] with applications to power grids [4, 6, 11, 12].

The system survivability terms are integrated into discrete models of cyber-physical infrastructures in various forms under Stackelberg game formulations [5]. A subclass of these models using the number of cyber and physical components that are attacked and reinforced as the main variables has been studied in [19]. These models characterize infrastructures with a large number of components, and are coarser compared to the models that consider the attacks and reinforcements of individual cyber and physical components. Under these formulations, various forms of correlation functions are used to capture the dependencies amongst the constituent systems and their components [16, 17, 19].

Collections of systems with complex interactions have been studied using game-theoretic formulations in [8], and their two-level correlations have been studied using the sum-form utility functions in [16] and the product-form utility functions in [17]. These two utility functions are unified in [13] and the sum-form utility function has been studied under the asymmetric role of communications network in [15]. These two works were unified in [14] by using the composite utility functions and additionally explicitly accounted for the asymmetric network role. The multi-site cloud computing infrastructure was discussed as an example for sum-form and product-form utility functions in [15] and composite utility function in [14] under the asymmetric role of the communications network. In this paper, we develop a comprehensive treatment of this infrastructure by including HVAC system and providing complete details of NE conditions and capacity estimates. In particular, we relate the abstract definitions of correlation functions and system multiplier functions to components of multi-site server infrastructure.

3 MULTI-SITE SERVER INFRASTRUCTURE

A distributed cloud computing infrastructure consisting of N sites, each with l_i servers at site i, $i = 1, 2, \ldots, N$ has been studied in [13] by using separate cyber and physical models for each site. The sites are connected over a communication network S_{N+1} as shown in Figure 1. The network consists of a number of routers each of which manages l_{N+1} connections as shown in Figure 2.

This infrastructure is subject to a variety of cyber and physical attacks on its components. Cyber attacks on the servers may be

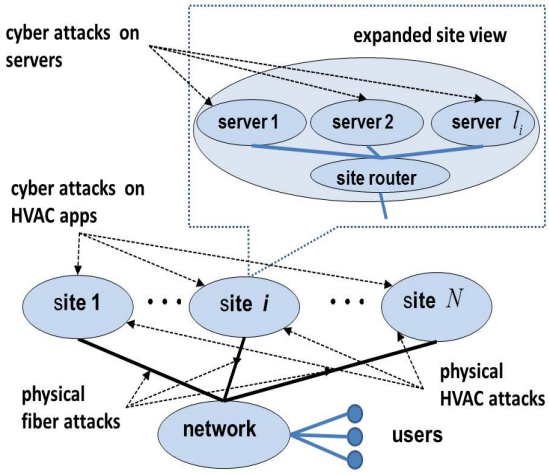

Figure 1: Cloud computing infrastructure with N server sites.

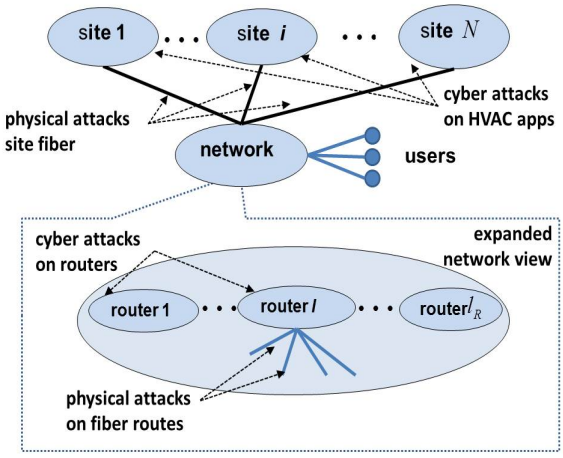

Figure 2: Network of multi-site cloud servers infrastructure.

launched remotely over the network since the servers are accessible to users. Meanwhile, routers are located at geographically separated sites and access to them is limited (to network administrators), and they are not as easily accessible over the network. Cyber attacks on routers require different techniques and represent different costs to the attacker compared to server attacks. Furthermore, this infrastructure is subject to physical attacks in the form of fiber cuts, which require a proximity access by the attacker. Cutting the network fibers that connect server sites to routers will disconnect the entire site, making it inaccessible to the users. And, such attacks may also be launched on the network fibers between routers at different locations on the network.

The infrastructure provider may employ a number of reinforcements to protect against attacks, including replicating the servers and routers to support fail-over operations, and installing physically separated redundant fiber lines to the sites and between router locations. These measures could require significant costs, and hence must be strategically chosen.

3.1 System-Level Correlations

The correlations between systems, including the network, in these infrastructure are characterized in terms of their survival probabilities as follows.

Condition 3.1. **Aggregate Correlation Function [16, 17]:** *Let C_i denote the failure probability of rest of the infrastructure S_{-i} given the failure of S_i, and let C_{-i} denote the failure probability of S_i given the failure of S_{-i}. Then, the survival probability of the infrastructure is given by*

$$\begin{aligned} P_I &= P_i + P_{-i} - 1 + C_i(1 - P_i) \\ &= P_i + P_{-i} - 1 + C_{-i}(1 - P_{-i}). \square \end{aligned}$$

The cyber and physical aspects of a site S_i can be represented by using two finer models $S_{(i,c)}$ and $S_{(i,p)}$ that correspond to cyber and physical model, respectively. Similarly, those of the network S_{N+1} are represented by $S_{(N+1,c)}$ and $S_{(N+1,p)}$, which are the cyber and physical models as illustrated in Figure 3. The relationships between these system-level models can be captured using refined versions of the aggregate correlation function defined above. For the communications network, we have

$$C_{(N+1,c)} = l_{N+1} C_{(N+1,p)}$$

which reflects that a cyber attack on a router will disrupt all its l_{N+1} connections, thereby illustrating the amplification effect of the cyber attacks. For the server sites, we have a similar effect due to physical fiber attacks denoted by label p_f reflected by

$$C_{(i,p_f)} = l_i C_{(i,c)}$$

which indicates that at site S_i the fiber disruption will disconnect all its l_i servers. Similarly, the cyber attach on site HVAC app denoted by label c_h leads to

$$C_{(i,c_h)} = l_i C_{(i,c)}$$

which indicates that at site S_i the HVAC disruption will affect all its l_i servers.

It is useful to examine interesting special cases of the aggregate correlation function. Under the statistical independence of system failures we have $C_i = 1 - P_{-i}$, where P_{-i} is the survival probability of S_{-i}, since the failure probability of S_{-i} is not dependent on P_i. Substituting in the above condition, we have $P_I = P_i P_{-i}$ as expected. Generalizations of this condition include two interesting cases: (a) If $C_i > 1 - P_{-i}$, the failures in S_{-i} are *positively correlated* to those in S_i, indicating that the conditional failure probability S_{-i} given the failure of S_i is higher than the failure probability of S_{-i}. (b) If $C_i < 1 - P_{-i}$, failures in S_{-i} are *negatively correlated* to latter failures.

The important asymmetric role of the communications network is characterized using the following condition.

Condition 3.2. **Asymmetric Network and Uncorrelated Systems Conditions [15]:** *The aggregated correlation functions of S_i, $i = 1, 2, \ldots, N+1$ satisfy the conditions: (i) for the network S_{N+1}, we have $C_{N+1} = 1$, and (ii) for the constituent systems, we have $C_i = 0$, $i = 1, 2, \ldots, N$.* $\square$

The part (i) leads to $P_I = P_{-(N+1)}$ which indicates the role of rest of infrastructure $S_{-(N+1)}$ without the network. The part (ii) leads to $P_I = P_i + P_{-i} - 1$, $i = 1, 2, \ldots, N$, which linearly depends on each of failure probabilities of the constituent system S_i and rest of infrastructure S_{-i}. It is important to note that although direct correlations between the failures of the sites are zero in part (ii) above, these failures are still correlated through the network, namely, each failure is individually correlated to the network, and the network failures are correlated to rest of the infrastructure, namely, the server sites.

At the system-level, the effects of reinforcements and attacks can be separated using the two following conditions:

(i) first condition, $\frac{\partial P_{-i}}{\partial x_i} \approx 0$ for $i = 1, 2, \ldots, N$, indicates that reinforcing the server site S_i does not directly impact the survival probability of other sites or network; and

(ii) second condition, $\frac{\partial P_i}{\partial x_j} \approx 0$ for $i = 1, 2, \ldots, N+1$, $j = 1, 2, \ldots, N$ and $j \neq i$, indicates that reinforcing server site S_j does not directly impact the survival probability of server site S_i.

While the reinforcements to individual server sites or network are not directly reflected in other systems, their failures may still be correlated due to the underlying system structures as reflected in their aggregated correlation functions. These system-level considerations for the provider are captured by the following condition which is obtained by differentiating P_I in Condition 3.1 with respect to x_i and ignoring the terms corresponding to parts (i) and (ii) above.

Condition 3.3. **De-Coupled Reinforcement Effects:** *For P_I in Condition 3.1, we have for $i = 1, 2, \ldots, N+1$,*

$$\frac{\partial P_I}{\partial x_i} \approx (1 - C_i)\frac{\partial P_i}{\partial x_i} + (1 - P_i)\frac{\partial C_i}{\partial x_i}$$

for the provider. □

The condition indicates that the increment in P_I due to change in the number of reinforced components x_i is the sum of the increment in individual system survival probability P_i weighted by "non-correlation" term $(1 - C_i)$ and increment in correlation C_i weighted by failure probability $1 - P_i$. In the cases where C_i is a constant, we note that $\frac{\partial C_i}{\partial x_i} = 0$, which is the case under both parts of Condition 3.2.

3.2 Component-Level Correlations

The survival probabilities for server sites and network satisfy the following differential condition that specifies the correlations at the component level within each site and network [16, 18].

Condition 3.4. **System Multiplier Functions:** *The survival probabilities P_i and P_{-i} of system S_i and S_{-i}, respectively, satisfy the following conditions: there exist* system multiplier functions Λ_i *and* Λ_{-i} *such that*

$$\frac{\partial P_i}{\partial x_i} = \Lambda_i(x_1, \ldots, x_N, y_1, \ldots, y_N)P_i$$
$$\frac{\partial P_{-i}}{\partial x_i} = \Lambda_{-i}(x_1, \ldots, x_N, y_1, \ldots, y_N)P_{-i}$$

for $i = 1, 2, \ldots, N+1$. □

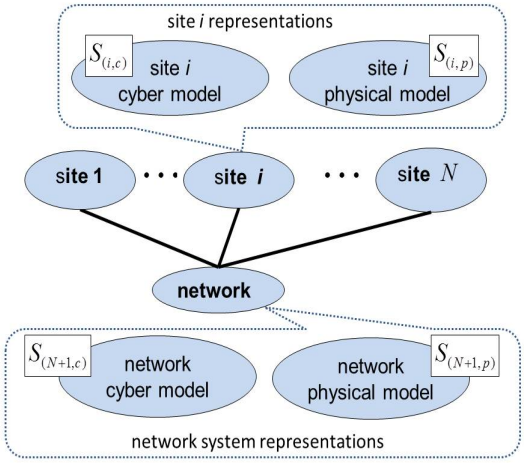

Figure 3: Representation of cloud computing infrastructure.

We now consider a special case where the attacker and provider choose the components to attack and reinforce, respectively, according to uniform distribution. Let $n_{(i,c)}$ and $n_{(i,p)}$ represent the number of cyber and physical components, respectively, of site S_i such that $n_i = n_{(i,c)} + n_{(i,p)}$. Similarly, let $x_{(i,c)}$ and $x_{(i,p)}$ represent the number of cyber and physical components reinforced at site S_i such that $x_i = x_{(i,c)} + x_{(i,p)}$, and let $y_{(i,c)}$ and $y_{(i,p)}$ represent the number of cyber and physical components attacked at site S_i such that $y_i = y_{(i,c)} + y_{(i,p)}$. Then, corresponding to the site physical model $S_{(i,p)}$, $i = 1, 2, \ldots, N$, there are $[n_{(i,p)} - x_{(i,p)}]_+$ non-reinforced fiber connections, where $[x]_+ = x$ for $x > 0$, and $[x]_+ = 0$ otherwise. Similarly, there are $[n_{(i,c)} - x_{(i,c)}]_+$ non-reinforced servers. If a cyber component (i.e., a server) is reinforced, it will survive a cyber attack but can be brought down indirectly by a fiber attack. Then, the probability that a cyber-reinforced component survives $y_{(i,p)}$ fiber attacks is approximated by

$$p_{(i,c)|R} = \frac{f_{(i,c)}}{1 + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+},$$

where the normalization constant $f_{(i,c)}$ is appropriately chosen.

On the other hand, if a cyber component is not reinforced, it can be brought down by either a direct cyber attack, or indirectly through a fiber attack. Thus, we approximate the survival probability of a cyber component at site k as

$$p_{(i,c)|N} = \frac{f_{(i,c)}}{1 + y_{(i,c)} + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+},$$

which reflects the additional lowering of the survival probability in inverse proportion to the level of cyber attack $y_{(i,c)}$. Using these formulae, for cyber model $S_{(i,c)}$ of site S_i, we have, under the independence of component attacks [18]

$$\Lambda_{(i,c)}(x_{(i,p)}, y_{(i,c)}, y_{(i,p)}) = \ln\left(1 + \frac{y_{(i,c)}}{1 + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+}\right).$$

It is interesting to note that the system multiplier function $\Lambda_{(i,c)}$ does not depend on the cyber reinforcements term $x_{(i,c)}$ even

though it corresponds to $\frac{\partial P_{(i,c)}}{\partial x_{(i,c)}}$. The function, however, depends on the physical reinforcement term $x_{(i,p)}$.

Under the statistical independence of cyber and physical attacks, we have the following generalization of the condition derived in [17]

$$P_i = p_{(i,c)|R}^{x_{(i,c)}} p_{(i,c)|N}^{n_{(i,c)}-x_{(i,c)}} p_{(i,p)|R}^{x_{(i,p)}} p_{(i,p)|N}^{n_{(i,p)}-x_{(i,p)}}$$

or equivalently

$$\ln P_i = n_{(i,c)} \ln p_{(i,c)|N} + x_{(i,c)} \ln \left(\frac{p_{(i,c)|R}}{p_{(i,c)|N}} \right) + n_{(i,p)} \ln p_{(i,p)|N} + x_{(i,p)} \ln \left(\frac{p_{(i,p)|R}}{p_{(i,p)|N}} \right)$$

By differentiating the equation with $x_{(i,c)}$, we obtain

$$\frac{\partial P_i}{\partial x_{(i,c)}} = \ln \left(\frac{p_{(i,c)|R}}{p_{(i,c)|N}} \right) P_i = \Lambda_{(i,c)} P_i.$$

Then, by noting that $\frac{\partial x_i}{\partial x_{(i,c)}} = 1$, we obtain

$$\frac{\partial P_i}{\partial x_i} = \Lambda_{(i,c)} P_i,$$

which enables us to approximate Λ_i by $\Lambda_{(i,c)}$.

4 NASH EQUILIBRIUM CONDITIONS

The provider's objective is to make the infrastructure resilient by reinforcing x_i components of S_i by optimizing the utility function. Similarly, the attacker's objective is to disrupt the infrastructure by attacking y_i components of S_i by optimizing the corresponding utility function. A game between an attacker and a provider involves balancing the costs of attacks and reinforcements of systems, given by $L_A(y_1, \ldots, y_{N+1})$ and $L_D(x_1, \ldots, x_{N+1})$, respectively, with the survival probability of the infrastructure.

4.1 Sum-Form and Product-Form Utility Functions

The *sum-form disutility function* is given by

$$U_{D+} = -[P_I(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1})]\, g_D + L_D(x_1, \ldots, x_{N+1}),$$

which will be minimized by the provider, and the scalar $g_D \geq 0$ represents the benefit of keeping the infrastructure operational. The *product-form disutility function* is given by

$$U_{D\times} = [1 - P_I(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1})] \times L_D(x_1, \ldots, x_{N+1}),$$

which will be minimized by the provider; it represents the "wasted" cost to the provider since it is the expected cost under the condition that the infrastructure fails. The sum-form and product-form utility functions [15] reflect two different values attached to keeping the infrastructure operational: the sum-form represents a weaker coupling of probability and cost terms, whereas the product-form utility function is their product. In general, they lead to qualitatively different defense strategies that are derived separately, and the corresponding expressions for the survival probabilities appear to be structurally different. The composite utility functions lead to simpler expressions for P_i, $i = 1, 2, \ldots, N$, and N_I at the Nash Equilibrium (NE). NE conditions are derived by equating the corresponding derivatives of the utility functions to zero, which yields the following for sum- and product-form utilities, respectively:

$$\frac{\partial U_{D+}}{\partial x_i} = \frac{\partial P_I}{\partial x_i} g_D - \frac{\partial C_D}{\partial x_i} = 0$$

$$\frac{\partial U_{D\times}}{\partial x_i} = -\frac{\partial P_I}{\partial x_i} C_D + (1 - P_I) \frac{\partial C_D}{\partial x_i} = 0$$

for $i = 1, 2, \ldots, N+1$ for the provider.

In particular, the dependence of P_i on cost terms and aggregate correlation functions, and their partial derivatives, can be presented in a compact form by using composite gain-cost and composite multiplier terms (to be defined in Section 4.2).

Under Conditions 3.1, 3.3, and 3.4, estimates of the survival probability of system S_i, for $i = 1, 2, \ldots, N+1$ are derived in [15]

$$\hat{P}_{i;D}^{A} = \frac{\frac{\partial C_i}{\partial x_i} - \xi_i^A}{\frac{\partial C_i}{\partial x_i} - (1 - C_i)\Lambda_i}$$

where $A = +$ and $A = \times$ correspond to sum-form and product-form, respectively, such that

$$\xi_i^A = \begin{cases} \frac{1}{g_D} \frac{\partial C_D}{\partial x_i} & \text{if } A = + \\ (1 - P_I) \frac{\partial \ln C_D}{\partial x_i}, & \text{if } A = \times \end{cases}$$

for $i = 1, 2, \ldots, N+1$ under the condition: $C_i < 1$ or $\frac{\partial C_i}{\partial x_i} \neq 0$. Under the asymmetric network correlation coefficient $C_{N+1} = 1$, the survival probability of the network is given by

$$P_{-(N+1);D}^{A} = \frac{\xi_{N+1}^A}{\Lambda_{-(N+1)}}$$

for $A = +, \times$.

In $\hat{P}_{i;D}^{A}$, the term Λ_i appears in the denominator with a negative sign. Thus, in qualitative terms, it depends linearly with a multiplier a on the logarithm of the number of cyber attacks $y_{(i,c)}$, and inversely on the logarithm of $\left[y_{(i,p)} - x_{(i,p)} \right]_+$ which is the number of attacks exceeding the reinforcements. The sign of the multiplier a could be positive or negative based on the other factors $\frac{\partial C_i}{\partial x_i}$ and ξ_i^A, where $A = +, \times$. This condition may appear somewhat counter-intuitive at the surface but note that it only characterizes the states that satisfy NE conditions, and in particular, it illustrates the richness of infrastructure behavior at NE.

4.2 Composite Utility Functions

The sum-form and product-form utility functions are generalized by the *composite utility function* given by

$$\begin{aligned} &U_D(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1}) \\ &= F_{D,G}(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1}) \\ &\quad \times G_D(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1}) \\ &\quad + F_{D,L}(x_1, \ldots, x_{N+1}, y_1, \ldots, y_{N+1}) L_D(x_1, \ldots, x_{N+1}), \end{aligned}$$

where the first product term corresponds to the reward and the second product term corresponds to the cost. Within the product terms, $F_{D,G}$ and $F_{D,L}$ are the reward and cost *multiplier functions*, respectively, of the provider, and G_D and L_D represent the reward

Table 1: Gain and cost terms and their multipliers for sum-form and product-form utilities of provider.

	$F_{D,G}$	G_D	$F_{D,L}$	L_D	$\frac{\partial F_{D,G}}{\partial P_I}$	$\frac{\partial G_D}{\partial x_i}$	$\frac{\partial F_{D,L}}{\partial P_I}$	$L^D_{G,L}$	$F^{D,i}_{G,L}$
sum-form: U_{D+}	$[1-P_I]$	g_D	1	L_D	- 1	0	0	$-g_D$	$\frac{\partial L_D}{\partial x_i}$
product-form: $U_{D\times}$	0	0	$[1-P_I]$	L_D	0	0	- 1	$-L_D$	$[1-P_I]\frac{\partial L_D}{\partial x_i}$

and cost, respectively, of keeping the infrastructure operational. Similarly, we consider that the attacker minimizes

$$\begin{aligned}&U_A(x_1,\ldots,x_{N+1},y_1,\ldots,y_{N+1})\\&= F_{A,G}(x_1,\ldots,x_{N+1},y_1,\ldots,y_{N+1})\\&\quad\times G_A(x_1,\ldots,x_{N+1},y_1,\ldots,y_{N+1})\\&\quad+ F_{A,L}(x_1,\ldots,x_{N+1},y_1,\ldots,y_{N+1})L_A(y_1,\ldots,y_{N+1}),\end{aligned}$$

where $F_{A,G}$ and $F_{A,L}$ are the reward and cost multiplier functions, respectively, of the attacker, and G_A and L_A represent the reward and cost of disrupting the infrastructure operation, respectively.

The provider's objective is to make the infrastructure resilient by reinforcing x_i components of S_i by optimizing the utility function. Similarly, the attacker's objective is to disrupt the infrastructure by attacking y_i components of S_i by optimizing the corresponding utility function. NE conditions are derived by equating the corresponding derivatives of the utility functions to zero, which yields

$$\frac{\partial U_D}{\partial x_i} = \left(G_D\frac{\partial F_{D,G}}{\partial P_I} + L_D\frac{\partial F_{D,L}}{\partial P_I}\right)\frac{\partial P_I}{\partial x_i} + F_{D,G}\frac{\partial G_D}{\partial x_i} + F_{D,L}\frac{\partial L_D}{\partial x_i} = 0$$

for $i = 1, 2, \ldots, N+1$ for the provider. We define

$$L^D_{G,L} = G_D\frac{\partial F_{D,G}}{\partial P_I} + L_D\frac{\partial F_{D,L}}{\partial P_I}$$

as the *composite gain-cost* term, wherein the gain G_D and loss L_D are "amplified" by the derivatives of their corresponding multiplier functions with respect to P_I. We then define

$$F^{D,i}_{G,L} = F_{D,G}\frac{\partial G_D}{\partial x_i} + F_{D,L}\frac{\partial L_D}{\partial x_i}$$

as the *composite multiplier* term, wherein the gain multiplier $F_{D,G}$ and cost multiplier $F_{D,L}$ are "amplified" by the derivatives of their corresponding gain and cost terms with respect to $x_i, i = 1, 2, \ldots, N+1$, respectively. These two terms lead to the compact NE condition $\frac{\partial P_I}{\partial x_i} = -\frac{F^{D,i}_{G,L}}{L^D_{G,L}}$. Various terms of the composite utility function specialized to sum-form and product-form utilities are shown in Table 1.

Under Conditions 3.1, 3.3, and 3.4, the following estimates of the survival probability of system S_i, for $i = 1, 2, \ldots, N+1$ are derived in [16]

$$\hat{P}_{i;D} = \frac{\frac{\partial C_i}{\partial x_i} + \frac{F^{D,i}_{G,L}}{L^D_{G,L}}}{\frac{\partial C_i}{\partial x_i} - (1-C_i)\Lambda_i}$$

for $i = 1, 2, \ldots, N+1$ under the condition: $C_i < 1$ or $\frac{\partial C_i}{\partial x_i} \neq 0$. Under the asymmetric network correlation coefficient $C_{N+1} = 1$, the survival probability of the network is given by

$$P_{-(N+1);D} = -\frac{1}{\Lambda_{-(N+1)}}\frac{F^{D,N+1}_{G,L}}{L^D_{G,L}}.$$

The system survival probability estimates $\hat{P}_{i;D}$ provide qualitative information about the effects of various parameters including aggregated correlation coefficient C_i, system multiplier functions Λ_i, composite gain-cost $L^D_{G,L}$ and composite multiplier $F^{D,i}_{G,L}$; note that the estimates may not necessarily lie within range [0,1]. In particular, $\hat{P}_{i;D}$ (i) increases and decreases with $F^{D,i}_{G,L}$ and $L^D_{G,L}$, respectively, (ii) increases with Λ_i, and (iii) depends both on C_i and its derivative for $i = 1, 2, \ldots, N$. For the network, $P_{-(N+1);D}$ is in a simpler form since $C_{N+1} = 1$.

Consider $\hat{P}^A_{i;D}$ above, the term Λ_i appears in the denominator with a negative sign. Thus, in qualitative terms, it depends linearly with a multiplier a on the logarithm of the number of cyber attacks $y_{(i,c)}$, and inversely on the logarithm of $\left[y_{(i,p)} - x_{(i,p)}\right]_+$ which is the number of attacks exceeding the reinforcements. The sign of the multiplier a could be positive or negative based on the other factors $\frac{\partial C_i}{\partial x_i}$ and ξ^A_i, where $A = +, \times$. This condition may appear somewhat counter-intuitive at the surface but note that it only characterizes the states that satisfy NE conditions, and in particular, it illustrates the richness of infrastructure behavior at NE.

5 EXPECTED CAPACITY ESTIMATES

We now consider that network failure renders the entire infrastructure unavailable, and those of individual systems are uncorrelated with others given by Condition 3.2.

5.1 Sum-Form and Product-Form Utility Functions

The following results derived in [15] provides a single, simplified expression for the expected capacity under these conditions. Under Conditions 3.1-3.4, the expected capacity is given by

$$N^A_I = \sum_{i=1}^{N}\left(n_i\frac{\xi^A_i}{\Lambda_i}\right)$$

where $A = +$ and $A = \times$ correspond to sum-form and product-form, respectively, such that

$$\xi^A_i = \begin{cases}\frac{1}{g_D}\frac{\partial C_D}{\partial x_i} & \text{if } A = +\\ (1-P_I)\frac{\partial \ln C_D}{\partial x_i}, & \text{if } A = \times\end{cases}$$

for $i = 1, 2, \ldots, N$. For the sum-form,

$$N^+_I = \sum_{i=1}^{N}\left(\frac{n_i\frac{\partial C_D}{\partial x_i}}{g_D\Lambda_i}\right)$$

indicates that higher gain g_D leads to lower number of operational components. For the product form,

$$N_I^{\times} = (1 - P_I) \sum_{i=1}^{N} \left(\frac{n_i \frac{\partial C_D}{\partial x_i}}{C_D \Lambda_i} \right)$$

indicates that higher survival probability of the network leads to lower number of operational components. The dependence on Λ_i is similar in both cases, namely, faster than linear leads to lower number of available component, and vice versa. The dependence on C_D is somewhat different due to its presence in the denominator for product-form, even though $\frac{\partial C_D}{\partial x_i}$ appears in the numerator in both forms.

In terms of the expected capacity N_I^A, the dependence on $y_{(i,c)}$ and $\left[y_{(i,p)} - x_{(i,p)}\right]_+$ is more direct, and qualitatively similar for both sum-form and product-form, since the term Λ_i appears in the denominator. We then obtain the following expressions: for the sum-form,

$$N_I^{+} = \sum_{i=1}^{N} \left(\frac{n_i \frac{\partial C_D}{\partial x_i}}{g_D \ln\left(1 + \frac{y_{(i,c)}}{1 + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+}\right)} \right),$$

and for the product form,

$$N_I^{\times} = (1 - P_I) \sum_{i=1}^{N} \left(\frac{n_i \frac{\partial C_D}{\partial x_i}}{C_D \ln\left(1 + \frac{y_{(i,c)}}{1 + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+}\right)} \right).$$

In both cases, the multipliers n_i, g_D and C_D are positive, and it is reasonable to assume the condition $\frac{\partial C_D}{\partial x_i} \geq 0$, since the reinforcement cost does not decrease with x_i. Thus, the expected capacity decreases with $y_{(i,c)}$ and the opposite is true with respect to $\left[y_{(i,p)} - x_{(i,p)}\right]_+$. In both cases, the dependence on the number of servers l_i at site i is qualitatively similar in that the expected capacity increases proportional to its logarithm. Thus, the overall dependencies considered here are quite simple, namely, under the statistical independence and uniform distributions of components chosen by both defender and attacker. Even under such simple conditions, the detailed NE conditions are quite complex to characterize.

For composite utility function, under Conditions 3.1-3.4, the expected capacity is derived in [14]

$$N_I = \sum_{i=1}^{N} \left(-\frac{n_i}{\Lambda_i} \frac{F_{G,L}^{D,i}}{L_{G,L}^{D}} \right)$$

for $i = 1, 2, \ldots, N$. This condition indicates that lower composite gain-cost $L_{G,L}^{D}$ and higher composite multiplier $F_{G,L}^{D,i}$ lead to lower expected capacity. Typically, the composite gain-cost $L_{G,L}^{D}$ is negative (e.g. $-g_D$ for sum-form) since it is minimized by the provider; thus, its lower value is more negative and has a higher magnitude. Also, larger values of Λ_i also lead to lower expected capacity. In particular, the condition $\Lambda_i > 1$, called the *faster than linear* growth of $\frac{\partial P_i}{\partial x_i}$, leads to lower expected capacity. This seems counter-intuitive since faster improvement in P_i due to increase in x_i leads to lower expected capacity, but note that it only characterizes the states that satisfy NE conditions.

For the composite utility functions, we obtain the following expression for the expected number of servers

$$N_I = \sum_{i=1}^{N} \left(-\frac{n_i F_{G,L}^{D,i}}{L_{G,L}^{D} \ln\left(1 + \frac{y_{(i,c)}}{1 + l_i \left[y_{(i,p)} - x_{(i,p)}\right]_+}\right)} \right).$$

In the equation, n_i is positive, and it is reasonable to assume that $-\frac{F_{G,L}^{D,i}}{L_{G,L}^{D}} \geq 0$, since $\frac{\partial P_I}{\partial x_i} = -\frac{F_{G,L}^{D,i}}{L_{G,L}^{D}}$ at NE, and the survival probability of entire infrastructure P_I does not decrease with x_i. Thus, the expected capacity decreases with $y_{(i,c)}$ and the opposite is true with respect to $\left[y_{(i,p)} - x_{(i,p)}\right]_+$. In both cases, the dependence on the number of servers l_i at site i is qualitatively similar in that the expected capacity increases proportional to its logarithm.

The dependencies considered here are quite simple as a result of the statistical independence and uniform distributions of reinforcements and attacks. Even under such simple conditions, the detailed NE conditions are quite complex to characterize, but they do provide qualitative insights into the effects of underlying parameters.

6 CONCLUSIONS

We consider a class of infrastructures with multiple server sites connected over a wide-area network, which plays an asymmetric role by providing the critical connectivity between them. By utilizing correlations at system- and component-level, we formulated the problem of ensuring the infrastructure survival as a game between an attacker and a provider, by using sum-form and product-form utility functions and their generalization using composite utility functions. We derived Nash Equilibrium conditions that provide compact expressions for the expected capacity given by the number of operational and accessible servers. These results are obtained by applying the extensions of previous results on interconnected systems [7, 8] and cyber-physical infrastructures [18] to the multi-site server infrastructure.

The formulation studied in this paper can be extended to include cases where targeted attacks and reinforcements of specific individual components are explicitly represented. It is of future interest to compare this formulation to ones whose utility functions explicitly utilize the capacity term in place of infrastructure survival probability terms. Another future direction is to consider the simultaneous cyber and physical attacks on multiple systems and components, and sequential game formulations of this problem. Performance studies of our approach using more detailed models of cloud computing infrastructure would be of future interest.

Acknowledgments

This work is funded by the Mathematics of Complex, Distributed, Interconnected Systems Program, Office of Advanced Computing Research, U.S. Department of Energy, and by Extreme Scale Systems Center, sponsored by U. S. Department of Defense, and performed at Oak Ridge National Laboratory managed by UT-Battelle, LLC for U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

REFERENCES

[1] V. M. Bier and M. N. Azaiez, editors. *Game Theoretic Risk Analysis of Security Threats*. Springer, 2009.

[2] S. Bu and F. R. Yu. A game-theoretical scheme in the smart grid with demand-side management: Towards a smart cyber-physical power infrastructure. *Emerging Topics in Computing, IEEE Transactions on*, 1(1):22–32, 2013.

[3] A. A. Cardenas, S. Amin, and S. Sastry. Secure control: Towards survivable cyber-physical systems. In *The 28th International Conference on Distributed Computing Systems Workshops*, pages 495–500. IEEE, 2008.

[4] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen. Smart attacks in smart grid communication networks. *Communications Magazine, IEEE*, 50(8):24–29, 2012.

[5] S. K. Das, K. Kant, and N. Zhang, editors. *An analytical framework for cyber-physical networks*. Morgan Kaufman, 2012.

[6] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu. Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid. *Smart Grid, IEEE Transactions on*, 4(2):847–855, 2013.

[7] K. Hausken. Strategic defense and attack of complex and dependent systems. *Reliability Engineering*, 95(1):29–42, 2009.

[8] K. Hausken. Defense and attack for interdependent systems. 2016.

[9] K. Hausken and G. Levitin. Review of systems defense and attack models. *International Journal of Performability Engineering*, 8(4):355–366, 2012.

[10] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)*, 45(3):25, 2013.

[11] Y. Mo, T. H.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli. Cyber–physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 100(1):195–209, 2012.

[12] F. Pasqualetti, F. Dörfler, and F. Bullo. Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 2195–2201. IEEE, 2011.

[13] N. S. V. Rao, N. Imam, C. Y. T. Ma, K. Hausken, F. He, and J. Zhuang. On defense strategies for system of systems using aggregated correlations. In *11th Annual IEEE International Systems Conference*, 2017.

[14] N. S. V. Rao, C. Y. T. Ma, K. Hausken, F. He, D. K. Y. Yau, and J. Zhuang. Defense strategies for asymmetric networked systems under composite utilities. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2017.

[15] N. S. V. Rao, C. Y. T. Ma, K. Hausken, F. He, D. K. Y. Yau, and J. Zhuang. Game-theoretic strategies for asymmetric networked systems. In *International Conference on Information Fusion*, 2017.

[16] N. S. V. Rao, C. Y. T. Ma, K. Hausken, F. He, and J. Zhuang. Defense strategies for infrastructures with multiple systems of components. In *International Conference on Information Fusion*, 2016.

[17] N. S. V. Rao, C. Y. T. Ma, K. Hausken, F. He, and J. Zhuang. Game-theoretic strategies for systems of components using product-form utilities. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2016.

[18] N. S. V. Rao, C. Y. T. Ma, F. He, J. Zhuang, and D. K. Y. Yau. Cyber-physical correlations for infrastructure resilience: A game-theoretic approach. In *International Conference on Information Fusion*, 2014.

[19] N. S. V. Rao, C. Y. T. Ma, U. Shah, J. Zhuang, F. He, and D. K. Y. Yau. On resilience of cyber-physical infrastructures using discrete product-form games. In *International Conference on Information Fusion*, 2015.

[20] S. Shiva, S. Roy, and D. Dasgupta. Game theory for cyber security. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, page 34. ACM, 2010.