Combating Informational Denial-of-Service (IDoS) Attacks: Modeling and Mitigation of Attentional Human Vulnerability

Linan Huang(✉) and Quanyan Zhu

Department of Electrical and Computer Engineering, New York University, 2 MetroTech Center, Brooklyn, NY 11201, USA
{lh2328,qz494}@nyu.edu

Abstract. This work proposes a new class of proactive attacks called the Informational Denial-of-Service (IDoS) attacks that exploit the attentional human vulnerability. By generating a large volume of feints, IDoS attacks deplete the cognitive resources of human operators to prevent humans from identifying the real attacks hidden among feints. This work aims to formally define IDoS attacks, quantify their consequences, and develop human-assistive security technologies to mitigate the severity level and risks of IDoS attacks. To this end, we use the semi-Markov process to model the sequential arrivals of feints and real attacks with category labels attached in the associated alerts. The assistive technology strategically manages human attention by highlighting selective alerts periodically to prevent the distraction of other alerts. A data-driven approach is applied to evaluate human performance under different Attention Management (AM) strategies. Under a representative special case, we establish the computational equivalency between two dynamic programming representations to reduce the computation complexity and enable online learning with samples of reduced size and zero delays. A case study corroborates the effectiveness of the learning framework. The numerical results illustrate how AM strategies can alleviate the severity level and the risk of IDoS attacks. Furthermore, the results show that the minimum risk is achieved with a proper level of intentional inattention to alerts, which we refer to as the *law of rational risk-reduction inattention*.

Keywords: Human vulnerability · Alert fatigue · Cyber feint attack · Temporal-difference learning · Risk analysis · Attention management · Cognitive load

1 Introduction

Human is the weakest link in cybersecurity due to their innate vulnerabilities, including bounded rationality and limited attention. These human vulnerabilities are difficult to mitigate through short-term training, rules, and incentives. As a result, sophisticated

Q. Zhu—This work is partially supported by grants SES-1541164, ECCS-1847056, CNS-2027884, and BCS-2122060 from National Science Foundation (NSF), and grant W911NF-19-1-0041 from Army Research Office (ARO).

B. Bošanský et al. (Eds.): GameSec 2021, LNCS 13061, pp. 314–333, 2021.
https://doi.org/10.1007/978-3-030-90370-1_17

attacks, such as Advanced Persistent Threats (APTs) and supply-chain attacks, commonly exploit them to breach data and damage critical infrastructures. Attentional vulnerabilities have been exploited by adversaries to create visual blindspots or misperceptions that can lead to erroneous outcomes. One way to exploit the attentional vulnerabilities is to stealthily evade the attention of human users or operators as we have seen in many cases of social engineering and phishing attacks. It is a passive approach where the attacker does not change the attention patterns of the human operators and intends to exploit the inattention to evade the detection. In contrast, a proactive attacker can strategically influence attention patterns. For example, an attacker can overload the attention of human operators with a large volume of feints and hide real attacks among them [1]. This class of proactive attacks aims to increase the perceptual and cognitive load of human operators to delay defensive responses and reduce detection accuracy. We refer to this class of attacks as the Informational Denial-of-Service (IDoS) attacks.

IDoS is no stranger to us in this age of information explosion. We are commonly overloaded with terabytes of unprocessed data or manipulated information on online media. However, the targeted IDoS attacks on specific groups of people, e.g., security guards, operators at the nuclear power plant, and network administrators, can pose serious threats to lifeline infrastructures and systems. The attacker customizes attack strategies to targeted individuals or organizations to quickly and maximally deplete their human cognitive resources. As a result, common methods (e.g., set tiered alert priorities) to mitigate alert fatigue are insufficient under these targeted and intelligent attacks that generate massive feints strategically. There is a need to understand this phenomenon, quantify its consequence and risks, and develop new mitigation methods. In this work, we establish a probabilistic model to formalize the definition of IDoS attacks, evaluate their severity levels, and assess the induced cyber risks. The model captures the interaction among attackers, human operators, and assistive technologies as highlighted by the orange, green, and blue backgrounds, respectively, in Fig. 1.

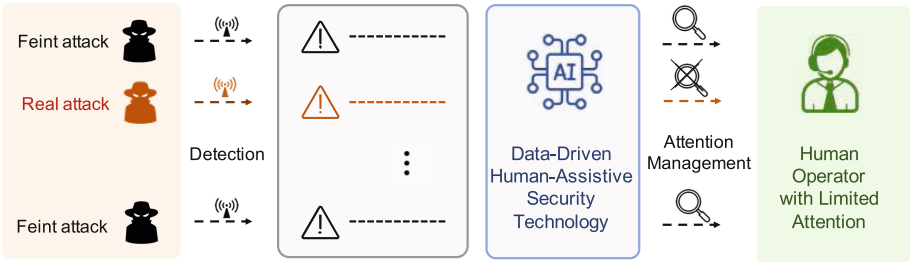


Fig. 1. Interaction among IDoS attacks, human operators, and assistive technologies.

Attackers generate feints and real attacks that trigger alerts of detection systems. Due to the detection imperfectness, human operators need to inspect these alerts in detail to determine the attacks' types, i.e., feint or real, and take responsive security decisions. The accuracy of the security decisions depends on the inspection time and the operator's sustained attention without distractions. The large volume of feints exerts

an additional cognitive load on each human operator and makes it hard to focus on each alert, which can significantly decrease the accuracy of his security decisions and increase cyber risks. Accepting the innate human vulnerability, we aim to develop assistive technologies to compensate for the human attention limitation. Evidence from the cognitive load theory [2] has shown that divided attention to multiple stimuli can degrade the performance and cost more time than responding to these stimuli in sequence. Hence, we design the *Attention Management* (AM) strategies to intentionally make some alerts inconspicuous so that the human operator can focus on the other alerts and finish the inspection with less time and higher accuracy. We further define risk measures to evaluate the inspection results, which serves as the stepping stone to designing adaptive AM strategies to mitigate attacks induced by human vulnerabilities.

Due to the unpredictability and complexity of human behaviors, cognition, and reasoning, it is challenging to create an exact human model of the IDoS attack response. Therefore, we provide a probabilistic characterization of human decisions concerning AM strategies and other observable features from the alerts. By assuming a sequential arrival of attacks with semi-Markov state transitions, we conduct a data-driven approach to evaluate the inspection results in real-time. Under a mild assumption, we prove the *computational equivalency* between two Dynamic Programming (DP) representations to simplify the value iteration and the Temporal-Difference (TD) learning process. Numerical results corroborate the effectiveness of learning by showing the convergence of the estimated value to the theoretical value. Without an AM strategy, we show that both the severity level and the risk of IDoS attacks increase with the product of the arrival rate and the detection threshold. With the assistance of AM strategies, we illustrate how different AM strategies can alleviate the severity level of IDoS attacks. Concerning the IDoS risks, we illustrate the tradeoff between the quantity and quality of the inspection, which leads to a meta-principle referred to as the *law of rational risk-reduction inattention*.

1.1 Related Works

Human Vulnerability in Cyber Space. Attacks that exploit human vulnerabilities, e.g., insider threats and social engineering, have raised increasing concerns in cybersecurity. Previous works have focused to design security rules [3] and incentives [4] to increase human employees' compliance and elicit desirable behaviors. However, compared to the lack of security awareness and incentives, some human vulnerabilities (e.g., attention limitation and bounded rationality) cannot be altered or controlled. Thus, we need to design assistive technologies to compensate for the 'unpatchable' human vulnerabilities. In [5], adaptive attention enhancement strategies have been developed to engage users' attention and maximize the rate of phishing recognition. Compared to [5] that defends against stealthy attacks and the exploitation of inattention, this work combats proactive attackers that overload human attention.

Data-Driven Approach for Security and Resilience. As more data becomes available, data-driven approaches have been widely used to create cyber situational awareness and enhance network security and resilience [6], e.g., Bayesian learning for parameter uncertainty [7,8] and Q-learning for honeypot engagement [9]. The authors in

[10] have studied the detection of feint attacks by a few-shot deep learning algorithm. However, they have modeled feints as multi-stage attacks and focused on detecting the revised causal relationship. Here, we focus on how feints affect human operators' cognitive resources and the consequent security decisions. The TD learning method helps address the long-standing challenge of human modeling and further enables us to evaluate human performance efficiently and robustly.

1.2 Notations and Organization of the Paper

We summarize notations in Table 1. The rest of the paper is organized as follows. Section 2 introduces the system modeling for IDoS attacks, alert generations, and the inspections of human operators. Based on the system model, we present a Semi-Markov Process (SMP) model in Sect. 3 to evaluate human performance, the severity level, and the risks of IDoS attacks. We present a case study in Sect. 4 to corroborate our results and Sect. 5 concludes the paper.

Table 1. Summary of variables and their meanings.

Variable	Meaning
$t^k \in [0,\infty)$	Arrival time of the k-th attack
$\tau^k = t^{k+1} - t^k \in [0,\infty)$	Time duration between k-th and $(k+1)$-th attack
$\tau_{IN}^{h,m} := \sum_{k'=hm}^{hm+m-1} \tau^{k'}$	Inspection time at inspection stage $h \in \mathbb{Z}^{0+}$
$w^k \in \mathscr{W} := \{w_{FE}, w_{RE}, w_{UN}\}$	Security decision at attack stages $k \in \mathbb{Z}^{0+}$
$a_m \in \mathscr{A}$	Attention management strategy of period $m \in \mathbb{Z}^{+}$
$\theta^k \in \Theta := \{\theta_{FE}, \theta_{RE}\}$	Attack's type at attack stages $k \in \mathbb{Z}^{0+}$
$\bar{\theta}^h := [\theta^{hm}, \cdots, \theta^{hm+m-1}]$	Consolidated type at inspection stage $h \in \mathbb{Z}^{0+}$
$s^k \in \mathscr{S}$	Alert's category label at attack stages $k \in \mathbb{Z}^{0+}$
$x^h := [s^{hm}, \cdots, s^{hm+m-1}]$	Consolidated state at inspection stage $h \in \mathbb{Z}^{0+}$
$\mathrm{Tr}(s^{k+1} \mid s^k; \theta^k)$	Transition probability from s^k to s^{k+1} under attack type θ^k
$\bar{T}r(x^{h+1} \mid x^h; \bar{\theta}^h)$	Transition function of the consolidated state

2 System Modeling of Informational Denial-of-Service Attacks

In Sect. 2.1, we present a high-level structure of the Informational Denial-of-Service (IDoS) attacks and use a motivating example to illustrate their causes, consequences, and mitigation methods. Then, we introduce the system modeling of sequential arrivals of alerts that are triggered by feints and real attacks in Sect. 2.2. The manual inspection and the attention management strategies are introduced in Sect. 2.3. Human operators inspect each alert in real-time to determine the associated attack's hidden type. Meanwhile, the assistive technology automatically designs and implements the optimal attention management strategy to compensate for human attention limitations.

2.1 High-Level Abstraction and Motivating Example

As shown in Fig. 2, there is an analogy between the Denial-of-Service (DoS) attacks in communication networks and the Informational Denial-of-Service (IDoS) attacks in the human-in-the-loop systems. Both of them achieve their attack goals by exhausting the limited resources. DoS attacks happen when the attacker generates a large number of superfluous requests to deplete the computing resource of the targeted machine and prevent the fulfillment of legitimate services. Analogously, IDoS attacks create a large amount of unprocessed information to deplete cognitive resources of human operators and prevent them from acquiring the knowledge contained in the information. We list

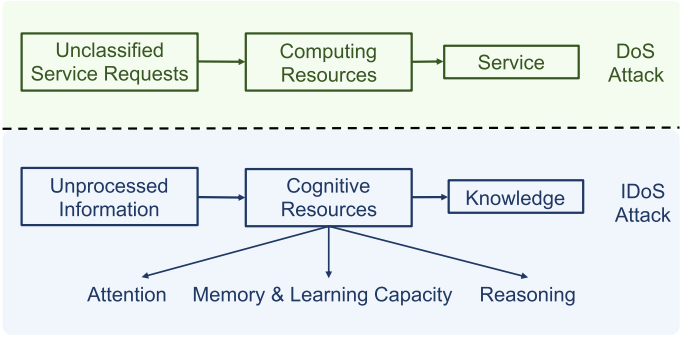


Fig. 2. The service request fulfillment process under DoS attacks and the information processing flows under IDoS attacks in green and blue backgrounds, respectively.

several assailable cognitive resources under IDoS attacks as follows.

- **Attention**: Paying sustained attention to acquire proper information is costly. From an economic perspective, inattention occurs when the cost of information acquisition is lower than the attention cost measured by the information entropy [11]. IDoS attacks generate feints to distract the human from the right information. An excessive number of feints prohibit the human from process any information.
- **Memory and Learning Capacity**: Humans have limited memory and learning capacity. Humans cannot remember the details or learn new things if there is an information overload [2].
- **Reasoning**: Human decision-making consumes a large amount of energy, which is one of the reasons why we have two modes of thought [12] ('system 1' thinking is fast, instinctive, and emotional; while 'system 2' thinking is slower and more logical). IDoS attacks can exert a heavy cognitive load to prevent humans from deliberative decisions that use the 'system 2' thinking. Moreover, evidence shows the *paradox of choice* [13]; i.e., rich choices can bring anxiety and prevent humans from making any decisions.

When these cognitive resources are exhausted, the information cannot be processed correctly and timely and serves as noise that leads to *alert fatigue* [14]. We use operators in the control room of nuclear power plants as a stylized example to illustrate

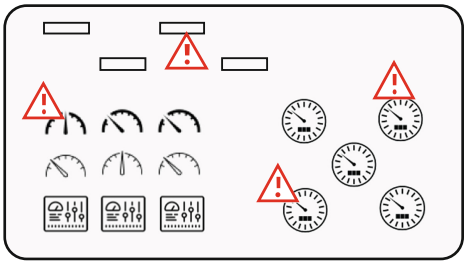


Fig. 3. A stylized example of the monitor screen for operators in the control room of nuclear power plants. The red triangles represent warnings and security messages.

the consequences of IDoS attacks and motivate the need for the security technology to assist human operators against IDoS attacks. In Fig. 3, a monitor screen contains meters that show the real-time readings of the temperature, pressure, and flow rate in a nuclear power plant. Based on the pre-defined generation rules, warnings and messages pop up at different locations. Due to the complexity of the nuclear control system, the inspection of these alerts consumes the operator's time and cognitive resources. The attempt to inspect all alerts and the constant switching among them can lead to missed detection and erroneous behaviors. If the alerts are generated strategically by attacks, they may further mislead humans to take actions in the attacker's favor; e.g., focusing on feints and ignoring the real attacks that hide among feints.

One way to mitigate IDoS attacks is to train the operators or human users to deal with the information overload and remain vigilant and productive under a heavy cognitive load. However, attentional training can be time-consuming and the effectiveness is not guaranteed. The second method is to recruit more human operators to share the information load. It would require the coordination of the operator team and can incur additional costs of human resources. The third method is to develop assistive technologies to rank and filter the information to alleviate the cognitive load of human operators. It would leverage past experiences and data analytics to pinpoint and prioritize critical alerts for human operators to process. The first two methods aim to increase the capacity or the volume of the cognitive resources in Fig. 2. The third method pre-processes the information so that it adapts to the capacity and characteristics of cognitive resources.

2.2 Sequential Arrivals of Alerts Triggered by Feints and Real Attacks

In this work, we focus on the temporal aspect of the alerts (i.e., the frequency and duration of their arrivals). The future work will incorporate their spatial locations on the monitor screen as shown in Fig. 3. As highlighted by the orange background in Fig. 4, attacks arrive sequentially at time $t^k, k \in \mathbb{Z}^{0+}$ where $t^0 = 0$. Let $\tau^k := t^{k+1} - t^k \in [0, \infty)$ be the inter-arrival time between the $(k+1)$-th attack and the k-th attack for all $k \in \mathbb{Z}^{0+}$. We refer to the k-th attack equivalently as the one at *attack stage* $k \in \mathbb{Z}^{0+}$.

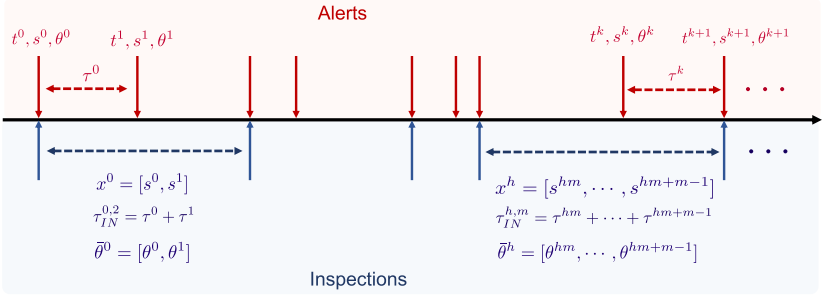


Fig. 4. The sequential arrival of alerts at *attack stage* $k \in \mathbb{Z}^{0+}$ and the periodic manual inspections at *inspection stage* $h \in \mathbb{Z}^{0+}$ under AM strategy $a_m \in \mathscr{A}$ where $m = 2$.

Each attack can be either a feint (denoted by θ_{FE}) or a real attack (denoted by θ_{RE}) with probability $b_{FE} \in [0,1]$ and $b_{RE} \in [0,1]$, respectively, where $b_{FE} + b_{RE} = 1$. We assume that both types of attacks trigger alerts with the same time delay. Thus, there is a one-to-one mapping between the sequence of attacks and alerts, and we can consider the zero delay time without loss of generality. The alerts cannot reflect the *attack's type* denoted by $\theta^k \in \Theta := \{\theta_{FE}, \theta_{RE}\}$ at all attack stages $k \in \mathbb{Z}^{0+}$. However, the alerts can provide human operators with a *category label* from a finite set $\mathscr{S}$ based on observable features or traces of the associated attacks, e.g., the attack locations as shown in Sect. 4. We denote the alert's category label at attack stage $k \in \mathbb{Z}^{0+}$ as $s^k \in \mathscr{S}$.

2.3 Manual Inspection and Attention Management

Since an alert does not directly reflect whether the attack is feint or real, human operators need to inspect the alert to determine the hidden type, which leads to three *security decisions*: the attack is feint (denoted by w_{FE}), the attack is real (denoted by w_{RE}), or the attack's type is unknown (denoted by w_{UN}). We use $w^k \in \mathscr{W} := \{w_{FE}, w_{RE}, w_{UN}\}$ to denote the human operator's security decision of the k-th alert. Each human operator has limited attention and cannot inspect multiple alerts simultaneously. Moreover, the human operator requires sustained attention on an alert to make an accurate security decision. Frequent alert pop-ups can distract humans from the current alert inspection and result in *alert fatigue* and the *paradox of choice* as illustrated in Sect. 2.1. To compensate for the human's attention limitation, we can intentionally make some alerts less noticeable, e.g., without sounds or in a light color. Then, the human can pay sustained attention to the alert currently under inspection. These inconspicuous alerts can be assigned to other available inspectors with an additional cost of human resources. If these alerts are time-insensitive, they can also be queued and inspected later by the same operator at his convenience. However, in practice, the number of alerts usually far exceeds the number of available inspectors, and the alerts cannot tolerate delay. Then, these alerts are dismissed as a tradeoff for the timely and accurate inspection of the other highlighted alerts. In this case, these inconspicuous alerts are not inspected and automatically assigned the security decision w_{UN}.

In this paper, we focus on the class of *Attention Management (AM) strategies*, denoted by $\mathscr{A} := \{a_m\}_{m\in\mathbb{Z}^+}$, that highlight alerts periodically to engage operators in the alert inspection. We assume that the human operator can only notice and inspect an alert when it is highlighted. Then, AM strategy $a_m \in \mathscr{A}$ means that the human operator inspects the alerts at attack stages $k = hm, h \in \mathbb{Z}^{0+}$. We refer to the attack stages during the h-th inspection as the *inspection stage* $h \in \mathbb{Z}^{0+}$. Then, under AM strategy $a_m \in \mathscr{A}$, each inspection stage contains m attack stages as shown in the blue background of Fig. 4. The h-th inspection has a duration of $\tau_{IN}^{h,m} := \sum_{k'=hm}^{hm+m-1} \tau^{k'}$ for all $h \in \mathbb{Z}^{0+}$.

Decision Probability with N Thresholds. The human operator's security decision depends on the attack's type, the category label, and the AM strategy. We refer to $\Pr(w^k|s^k, a_m; \theta^k)$ as the *decision probability*; i.e., the probability of human making decision $w^k \in \mathscr{W}$ when the attack's type is $\theta^k \in \Theta$, the category label is $s^k \in \mathscr{S}$, and the AM strategy is $a_m \in \mathscr{A}$. As a probability measure, the decision probability satisfies $\sum_{w^k\in\mathscr{W}} \Pr(w^k|s^k, a_m; \theta^k) = 1, \forall\theta^k \in \Theta, \forall s^k \in \mathscr{S}, \forall a_m \in \mathscr{A}$.

At attack stages where alerts are inconspicuous, i.e., for all $k \neq hm, h \in \mathbb{Z}^{0+}$, the security decision w^k is w_{UN} with probability 1; i.e., for any given inspection policy $a_m \in \mathscr{A}$, we have $\Pr(w^k|s^k, a_m; \theta^k) = \mathbf{1}_{\{w^k=w_{UN}\}}, \forall s^k \in \mathscr{S}, \forall w^k \in \mathscr{W}, \forall\theta^k \in \Theta, \forall k \neq hm, h \in \mathbb{Z}^{0+}$. At attack stages of highlighted alerts, i.e., for all $k = hm, h \in \mathbb{Z}^{0+}$, the human operator inspects the h-th alert for a duration of $\tau_{IN}^{h,m}$. At each inspection stage h, a longer period length m induces a longer inspection time $\tau_{IN}^{h,m} = \sum_{k'=hm}^{hm+m-1} \tau^{k'}$. Based on the IDoS model in Sect. 2, different AM strategies only affect the inspection time. Thus, we can rewrite the decision probability $\Pr(w^k|s^k, a_m; \theta^k)$ as $\Pr(w^k|s^k, \tau_{IN}^{h,m}; \theta^k)$ at attack stages $k = hm, h \in \mathbb{Z}^{0+}$.

Adequate inspection time $\tau_{IN}^{h,m}$ leads to an accurate security decision. In this work, we assume that the probability of correct decision-making can be approximated by an increasing step function of the inspection time as shown in Fig. 5. That is, $N+1$ thresholds divide the support of the random variable $\tau_{IN}^{h,m}$, i.e., $[0,\infty)$, into N regions where the probability of correct security decisions increases. We can increase the number of thresholds, i.e., the value of N, to improve the accuracy of the approximation. For each $s^k \in \mathscr{S}$ and $\theta^k \in \Theta$, we denote the corresponding N thresholds as $\bar{\tau}_n(s^k, \theta^k) \in \mathscr{N}(s^k, \theta^k), n \in \{0, 1, \cdots, N\}$, where $\mathscr{N}(s^k, \theta^k)$ is a finite set, $\bar{\tau}_0(s^k, \theta^k) = 0$, $\bar{\tau}_N(s^k, \theta^k) = \infty$, and $\bar{\tau}_0(s^k, \theta^k) < \bar{\tau}_N(s^k, \theta^k) < \bar{\tau}_2(s^k, \theta^k) < \cdots < \bar{\tau}_N(s^k, \theta^k)$. If $\tau_{IN}^{h,m}$ belongs to the region $n \in \{0, 1, \cdots, N\}$, i.e., $\bar{\tau}_{n-1}(s^k, \theta^k) < \tau_{IN}^{h,m} < \bar{\tau}_n(s^k, \theta^k)$, then the decision probabilities under θ_{FE} and θ_{RE} are represented as (1) and (2), respectively,

$$\Pr(w^k|s^k, \tau_{IN}^{h,m}; \theta_{FE}) = \begin{cases} \bar{p}_{CD}^{n-1}(s^k, \theta_{FE}) \in [0,1] & \text{if } w^k = w_{FE} \\ \bar{p}_{ID}^{n-1}(s^k, \theta_{FE}) \in [0,1] & \text{if } w^k = w_{RE} \\ 1 - \bar{p}_{CD}^{n-1}(s^k, \theta_{FE}) - \bar{p}_{ID}^{n-1}(s^k, \theta_{FE}) & \text{if } w^k = w_{UN} \end{cases} \quad (1)$$

and

$$\Pr(w^k|s^k, \tau_{IN}^{h,m}; \theta_{RE}) = \begin{cases} \bar{p}_{CD}^{n-1}(s^k, \theta_{RE}) \in [0,1] & \text{if } w^k = w_{RE} \\ \bar{p}_{ID}^{n-1}(s^k, \theta_{RE}) \in [0,1] & \text{if } w^k = w_{FE} \\ 1 - \bar{p}_{CD}^{n-1}(s^k, \theta_{RE}) - \bar{p}_{ID}^{n-1}(s^k, \theta_{RE}) & \text{if } w^k = w_{UN} \end{cases} \quad (2)$$

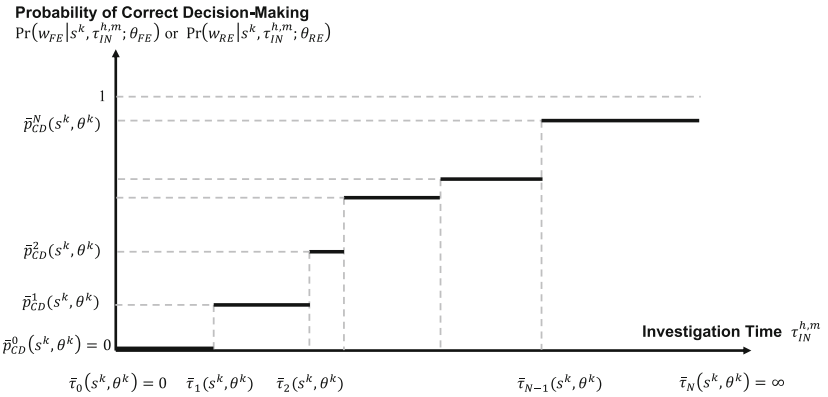


Fig. 5. The probability of the human operator making correct security decisions, i.e., $\Pr(w_{FE}|s^k,\tau_{IN}^{h,m};\theta_{FE})$ and $\Pr(w_{RE}|s^k,\tau_{IN}^{h,m};\theta_{RE})$, is approximated as an increasing step function of the inspection time $\tau_{IN}^{h,m}$ at inspection stage $h\in\mathbb{Z}^{0+}$.

In both (1) and (2), the first and second cases represent the probability of making correct and incorrect security decisions, respectively. The third case represents the probability that the human operator is uncertain about the attack's type and needs more time to inspect. A longer inspection time has two impacts:

- Increases the probability of making correct security decisions, i.e., $0=\bar{p}_{CD}^{0}(s^k,\theta^k)\leq\bar{p}_{CD}^{1}(s^k,\theta^k)\leq\cdots\leq\bar{p}_{CD}^{N}(s^k,\theta^k)\leq 1$, for any given $s^k\in\mathscr{S}$ and $\theta^k\in\Theta$.
- Decreases the probability of incorrect security decisions, i.e., $0\leq\bar{p}_{ID}^{N}(s^k,\theta^k)\leq\bar{p}_{ID}^{N-1}(s^k,\theta^k)\leq\cdots\leq\bar{p}_{ID}^{0}(s^k,\theta^k)\leq 1$, for any given $s^k\in\mathscr{S}$ and $\theta^k\in\Theta$.

3 Semi-Markov Process Model for Performance Evaluation

We assume that the category label of the sequential attacks follows a semi-Markov process based on the attack's type where $\mathrm{Tr}(s^{k+1}|s^k;\theta^k)$ represents the *transition probability* from $s^k\in\mathscr{S}$ to $s^{k+1}\in\mathscr{S}$ when the attack's type is $\theta^k\in\Theta$ at attack stage $k\in\mathbb{Z}^{0+}$. As a probability measure, the transition probability satisfies $\sum_{s^{k+1}\in\mathscr{S}}\mathrm{Tr}(s^{k+1}|s^k;\theta^k)=1,\forall s^k\in\mathscr{S},\forall\theta^k\in\Theta$. The inter-arrival time τ^k is a continuous random variable with a Probability Density Function (PDF) denoted by $z(\cdot|s^k;\theta^k)$.

3.1 Consolidated State and Consolidated Cost

Since the inspection is made every m attack stages, we define the *consolidated state* $x^h:=[s^{hm},\cdots,s^{hm+m-1}]\in\mathscr{X}:=\mathscr{S}^m$ that consists of the category labels of m successive alerts at inspection stage $h\in\mathbb{Z}^{0+}$. Analogously, we define the *consolidated type* $\bar{\theta}^h:=[\theta^{hm},\cdots,\theta^{hm+m-1}]\in\bar{\Theta}:=\Theta^m$. Then, we denote the transition function of the consolidated state as $\bar{T}r(x^{h+1}|x^h;\bar{\theta}^h)$, which is also Markov as shown below.

$$\begin{aligned}
&\Pr(x^{h+1}|x^h,\cdots,x^1;\bar{\theta}^h,\cdots,\bar{\theta}^1) = \frac{\Pr(x^{h+1},x^h,\cdots,x^1;\bar{\theta}^h,\cdots,\bar{\theta}^1)}{\Pr(x^h,\cdots,x^1;\bar{\theta}^h,\cdots,\bar{\theta}^1)} \\
&= \frac{\Pr(s^{(h+2)m-1}|s^{(h+2)m-2};\theta^{(h+2)m-2})\Pr(s^{(h+2)m-2}|s^{(h+2)m-3};\theta^{(h+2)m-3})\cdots\Pr(s^1\,s^0;\theta^0)}{\Pr(s^{(h+1)m-1}|s^{(h+1)m-2};\theta^{(h+1)m-2})\Pr(s^{(h+1)m-2}|s^{(h+1)m-3};\theta^{(h+1)m-3})\cdots\Pr(s^1\,s^0;\theta^0)} \\
&= \Pr(s^{(h+2)m-1}|s^{(h+2)m-2};\theta^{(h+2)m-2})\cdots\Pr(s^{(h+1)m-1}|s^{(h+1)m-2};\theta^{(h+1)m-2}) \\
&= \bar{T}r(x^{h+1}|x^h;\bar{\theta}^h).
\end{aligned} \tag{3}$$

The *inspection time* $\tau_{IN}^{h,m} = \sum_{k'=hm}^{hm+m-1} \tau^{k'}$ at inspection stage $h \in \mathbb{Z}^{0+}$ is a continuous random variable with support $[0,\infty)$ whose PDF $\bar{z}(\cdot|x^h;\bar{\theta}^h)$ can be computed based on the PDF z. Based on $\bar{z}$ and $\Pr(w^{hm}|s^{hm},\tau_{IN}^{h,m};\theta^{hm})$ in (1) and (2), we can compute the probability of security decision w^{hm} at inspection stage $h \in \mathbb{Z}^{0+}$ given x^h and $\bar{\theta}^h$, i.e.,

$$\begin{aligned}
\Pr(w^{hm}|x^h,a_m;\bar{\theta}^h) &= \int_0^\infty \Pr(w^{hm},\tau_{IN}^{h,m}|x^h;\bar{\theta}^h)d(\tau_{IN}^{h,m}) \\
&= \int_0^\infty \Pr(w^{hm}|s^{hm},\tau_{IN}^{h,m};\theta^{hm})\bar{z}(\tau_{IN}^{h,m}|x^h,\bar{\theta}^h)d(\tau_{IN}^{h,m}).
\end{aligned} \tag{4}$$

Let $\Pr(w^{hm}|x^h,a_m;\theta^{hm})$ be the shorthand notation for $\mathbb{E}_{\theta^l\sim[b_{FE},b_{RE}],l\in\{hm+1,\cdots,hm+m-1\}}[\Pr(w^{hm}|x^h,a_m;\bar{\theta}^h)], \forall\theta^{hm}\in\Theta$. We define the probability of the human operator making correct security decisions at inspection stage $h \in \mathbb{Z}^{0+}$ as

$$\hat{p}_{CD}(x^h,a_m) := b_{FE}\Pr(w_{FE}|x^h,a_m;\theta_{FE}) + b_{RE}\Pr(w_{RE}|x^h,a_m;\theta_{RE}), \forall x^h \in \mathscr{X}, \tag{5}$$

which leads to the *consolidated severity level* of IDoS attacks in Definition 1.

Definition 1 (Consolidated Severity Level). *We define* $1-\hat{p}_{CD}(x^h,a_m)$ *as the consolidated severity level of IDoS attacks under the consolidated state* $x^h \in \mathscr{X}$ *and AM strategy* $a_m \in \mathscr{A}$.

We denote $c(w^k,s^k;\theta^k)$ as the operator's cost at attack stage $k \in \mathbb{Z}^{0+}$ when the alert's category label is $s^k \in \mathscr{S}$, the attack's type is $\theta^k \in \Theta$, and the security decision is $w^k \in \mathscr{W}$. At attack stages where alerts are inconspicuous, i.e., for all $k \neq hm, h \in \mathbb{Z}^{0+}$, the security decision is w_{UN} without manual inspection, which incurs an *uncertainty cost* $c_{UN} > 0$. At attack stages of highlighted alerts, i.e., for all $k = hm, h \in \mathbb{Z}^{0+}$, the human operator obtains a reward (resp. cost), denoted by $c_{CD}(s^k;\theta^k) < 0$ (resp. $c_{ID}(s^k;\theta^k) > 0$), for correct (resp. incorrect) security decisions. If the human operator remains uncertain about the attack's type after the inspection time $\tau_{IN}^{h,m}$, i.e., $w^{hm} = w_{UN}$, there is the uncertainty cost c_{UN}. We define the human operator's *consolidated cost* at inspection stage $h \in \mathbb{Z}^{0+}$ as

$$\bar{c}(x^h,a_m;\bar{\theta}^h) := (m-1)c_{UN} + \sum_{w^{hm}\in\mathscr{W}} \Pr(w^{hm}|x^h,a_m;\bar{\theta}^h)c(w^{hm},s^{hm};\theta^{hm}). \tag{6}$$

3.2 Long-Term Risk Measures for IDoS Attacks

In this section, we define four long-term risk measures whose relations are shown in Fig. 6. The *Cumulative Cost (CC)* and *Expected Cumulative Cost (ECC)* on the left

directly follow from the discounted summation of the consolidated cost $\bar{c}$ in (6). Since CC and ECC depend on the consolidated state x^h and the consolidated type $\bar{\theta}^h$, it is of high dimension and thus difficult to store and compute. By taking an expectation over $s^{hm+1}, \cdots, s^{hm+m-1} \in \mathscr{S}$, we reduce the dimension and obtain the Aggregated Cumulative Cost (ACC) and Expected Aggregated Cumulative Cost (EACC) on the right of the figure. The DP representations for CC (resp. ECC) and ACC (resp. EACC) are generally not equivalent. We identify the condition under which two DP representations are equivalent in Sect. 3.3. The two risk learning schemes are introduced in Sect. 3.4. Since the consolidated risk learning is based on ECC, it has to wait for the realization of the consolidated state $x^h := [s^{hm}, \cdots, s^{hm+m-1}]$ to evaluate the inspection performance. On the contrary, the EACC-based aggregated risk learning just needs s^{hm} to evaluate the inspection performance, which reduces the dimension of the samples and enables evaluations with no delay.

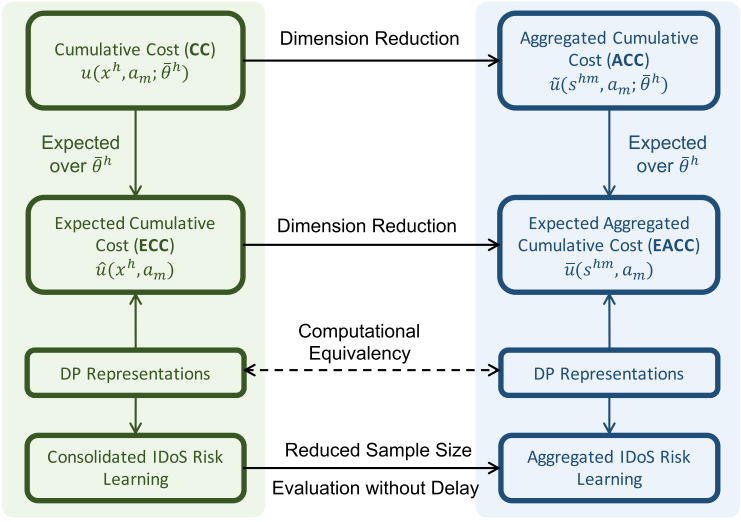


Fig. 6. Relations among four long-term risk measures, their DP representations, and two risk learning schemes.

Cumulative Cost and Expected Cumulative Cost. With discounted factor $\gamma \in (0,1)$, we define the *Cumulative Cost (CC)* under x^{h_0}, $\bar{\theta}^{h_0}$, and action a_m as $u(x^{h_0}, a_m; \bar{\theta}^{h_0}) := \mathbb{E}[\sum_{h=h_0}^{\infty} (\gamma)^h \cdot \bar{c}(x^h, a_m; \bar{\theta}^h)]$, where the expectation is taken over x^{h_0+n} and $\theta^{h_0 m+n}$ for all $n \in \{1, 2, \cdots, \infty\}$. By Dynamic Programming (DP), we represent u in the following iterative form, i.e., for all $x^h \in \mathscr{X}$, $\bar{\theta}^h \in \bar{\Theta}$, and $h \in \mathbb{Z}^{0+}$,

$$u(x^h, a_m; \bar{\theta}^h) = \bar{c}(x^h, a_m; \bar{\theta}^h) + \gamma \sum_{x^{h+1} \in \mathscr{X}} \bar{T}r(x^{h+1} | x^h; \bar{\theta}^h) \mathbb{E}_{\bar{\theta}^{h+1}}[u(x^{h+1}, a_m; \bar{\theta}^{h+1})]. \quad (7)$$

Denote $u^l(x^h, a_m; \bar{\theta}^h), l \in \mathbb{Z}^{0+}$, as the estimated value of $u(x^h, a_m; \bar{\theta}^h)$ at the l-th iteration, we can compute (7) by the following value iteration algorithm in Algorithm 1. It

can be shown that $u^{\infty}(x^h, a_m; \bar{\theta}^h)$ converges to $u(x^h, a_m; \bar{\theta}^h)$ and the following lemma holds [15].

Algorithm 1: Value Iteration

Initialize a stopping threshold $\varepsilon > 0$, $l = 0$, and $u^0(x^h, a_m; \bar{\theta}^h) = 0, \forall x^h \in \mathscr{X}, \bar{\theta}^h \in \bar{\Theta}$;
while $\max_{x^h \in \mathscr{X}, \bar{\theta}^h \in \bar{\Theta}} [u^{l+1}(x^h, a_m; \bar{\theta}^h) - u^l(x^h, a_m; \bar{\theta}^h)] \geq \varepsilon$ **do**
  for $x^h \in \mathscr{X}$ *and* $\bar{\theta}^h \in \bar{\Theta}$ **do**
    Update estimated value $u^{l+1}(x^h, a_m; \bar{\theta}^h) =$
    $\bar{c}(x^h, a_m; \bar{\theta}^h) + \gamma \sum_{x^{h+1} \in \mathscr{X}} \bar{T}r(x^{h+1}|x^h; \bar{\theta}^h) \mathbb{E}_{\bar{\theta}^{h+1}}[u^l(x^{h+1}, a_m; \bar{\theta}^{h+1})]$;
  end
  $l \leftarrow l + 1$;
end
Return $u^{l+1}(x^h, a_m; \bar{\theta}^h)$;

Lemma 1 (Monotonicity Lemma). *Let* $u'(x^{h_0}, a_m; \bar{\theta}^{h_0}) := \mathbb{E}[\sum_{h=h_0}^{\infty} (\gamma)^h \cdot \bar{c}'(x^h, a_m; \bar{\theta}^h)]$. *If* $\bar{c}(x^h, a_m; \bar{\theta}^h) > \bar{c}'(x^h, a_m; \bar{\theta}^h), \forall x^h \in \mathscr{X}, \bar{\theta}^h \in \bar{\Theta}$, *then* $u(x^h, a_m; \bar{\theta}^h) > u'(x^h, a_m; \bar{\theta}^h)$ *for all* $x^h \in \mathscr{X}, \bar{\theta}^h \in \bar{\Theta}$.

We define the *Expected Cumulative Cost (ECC)* as $\hat{u}(x^h, a_m) := \mathbb{E}_{\bar{\theta}^h}[u(x^h, a_m; \bar{\theta}^h)]$, $\forall x^h \in \mathscr{X}$, and write the DP representation of $\hat{u}$ in (8) by taking expectation over $\bar{\theta}^h$ in (7).

$$\hat{u}(x^h, a_m) = \mathbb{E}_{\bar{\theta}^h}[\bar{c}(x^h, a_m; \bar{\theta}^h)] + \gamma \sum_{x^{h+1} \in \mathscr{X}} \mathbb{E}_{\bar{\theta}^h}[\bar{T}r(x^{h+1}|x^h; \bar{\theta}^h)] \hat{u}(x^{h+1}, a_m). \tag{8}$$

Aggregated Cumulative Cost and Expected Aggregated Cumulative Cost. We define the *Aggregated Cumulative Cost (ACC)* as

$$\tilde{u}(s^{hm}, a_m; \bar{\theta}^h) := \sum_{s^{hm+1}, \cdots, s^{hm+m-1} \in \mathscr{S}} \Big[\Pr(s^{hm+1}, \cdots, s^{hm+m-1} | s^{hm}; \bar{\theta}^h) \cdot u([s^{hm}, \cdots, s^{hm+m-1}], a_m; \bar{\theta}^h) \Big], \tag{9}$$

and the *Expected Aggregated Cumulative Cost (EACC)* as

$$\bar{u}(s^{hm}, a_m) := \mathbb{E}_{\theta^l \sim [b_{FE}, b_{RE}], l \in \{hm, \cdots, hm+m-1\}} [\tilde{u}(s^{hm}, a_m; \bar{\theta}^h)], \forall s^{hm} \in \mathscr{S}. \tag{10}$$

Both ECC $\hat{u}(x^0, a_m)$ and EACC $\bar{u}^0(s^0, a_m)$ evaluate the long-term performance of the AM strategy $a_m \in \mathscr{A}$ on average as defined in Definition 2. However, EACC depends on s^{hm} but not on $s^{hm+1}, \cdots, s^{hm+m-1}$.

Definition 2 (Consolidated and Aggregated IDoS risks). *We define ECC* $\hat{u}(x^h, a_m)$ *(resp. EACC* $\bar{u}(s^{hm}, a_m)$*) as the consolidated (resp. aggregated) risk of the IDoS attack under* $x^h \in \mathscr{X}$ *(resp.* $s^{hm} \in \mathscr{S}$*) and attention strategy* $a_m \in \mathscr{A}$.

3.3 Inter-arrival Time with Independent PDF

In Sect. 3.3, we consider the special case where PDF z is independent of s^k and θ^k, which reduces the dependency of $\hat{p}_{CD}$ and $\bar{c}$ from x^h to s^{hm} as shown in Lemma 2. Moreover, we can obtain DP representations for ACC $\tilde{u}$ and EACC $\bar{u}$ as shown in Theorem 1. Value iteration in Algorithm 1 can be revised accordingly to solve these two DP representations.

Lemma 2. *If PDF z is independent of s^k and θ^k, then $\hat{p}_{CD}(x^h, a_m)$ in (5) can be rewritten as $\hat{p}_{CD}(s^{hm}, a_m)$ and the consolidated cost $\bar{c}(x^h, a_m; \bar{\theta}^h)$ in (6) can be rewritten as $\bar{c}(s^{hm}, a_m; \theta^{hm})$ without loss of generality.*

Proof. If z is independent of s^k, θ^k, then $\bar{z}$ is independent of $x^h, \bar{\theta}^h$, and $\Pr(w^{hm}|x^h, a_m; \bar{\theta}^h)$ in (4) only depends on s^{hm} and θ^{hm}. Thus, $\hat{p}_{CD}$ becomes a function of s^{hm}, a_m, and the consolidated cost $\bar{c}$ in (6) becomes a function of s^{hm}, θ^{hm}, and a_m. □

Theorem 1. *If PDF z is independent of s^k and θ^k, then we have the following DP representation in (11) for the ACC*

$$\tilde{u}(s^{hm}, a_m; \bar{\theta}^h) = \bar{c}(s^{hm}, a_m; \theta^{hm}) + \gamma \sum_{s^{(h+1)m} \in \mathscr{X}} \Pr(s^{(h+1)m}|s^{hm}; \bar{\theta}^h) \mathbb{E}_{\bar{\theta}^{h+1}}[\tilde{u}(s^{(h+1)m}, a_m; \bar{\theta}^{h+1})], \tag{11}$$

and the following DP representation in (12) for the EACC

$$\bar{u}(s^{hm}, a_m) = \mathbb{E}_{\theta^{hm}}[\bar{c}(s^{hm}, a_m; \theta^{hm})] + \gamma \sum_{s^{(h+1)m} \in \mathscr{X}} \mathbb{E}_{\bar{\theta}^h}[\Pr(s^{(h+1)m}|s^{hm}; \bar{\theta}^h)] \cdot \bar{u}(s^{(h+1)m}, a_m), \tag{12}$$

where

$$\mathbb{E}_{\bar{\theta}^h}[\Pr(s^{(h+1)m}|s^{hm}; \bar{\theta}^h)] = \sum_{s^{hm+1}, \cdots, s^{hm+m-1} \in \mathscr{S}} \prod_{l=hm}^{(h+1)m} \mathbb{E}_{\theta^l \sim [b_{FE}, b_{RE}]}[Tr(s^{l+1}|s^l; \theta^l)]. \tag{13}$$

Proof. First, for all $\bar{\theta}^h \in \bar{\Theta}$, we have

$$\sum_{s^{hm+1}, \cdots, s^{hm+m-1} \in \mathscr{S}} \Pr(s^{hm+1}, \cdots, s^{hm+m-1}|s^{hm}; \bar{\theta}^h) \cdot \bar{c}(s^{hm}, a_m; \theta^{hm}) \equiv \bar{c}(s^{hm}, a_m; \theta^{hm}).$$

Second, since

$$\begin{aligned} \bar{T}r(x^{h+1}|x^h; \bar{\theta}^h) &= \Pr(s^{(h+1)m}, \cdots, s^{(h+2)m-1}|s^{hm+m-1}; \bar{\theta}^h) \\ &= \Pr(s^{(h+1)m+1}, \cdots, s^{(h+2)m-1}|s^{(h+1)m}; \bar{\theta}^h) \mathrm{Tr}(s^{(h+1)m}|s^{hm+m-1}; \theta^{hm+m-1}) \end{aligned}$$

as shown in (3), we have

$$\begin{aligned}
&\sum_{s^{hm+1},\cdots,s^{hm+m-1}\in\mathscr{S}} \Pr(s^{hm+1},\cdots,s^{hm+m-1}|s^{hm};\bar{\theta}^h) \\
&\quad\cdot \sum_{x^{h+1}\in\mathscr{X}} \bar{T}r(x^{h+1}|x^h;\bar{\theta}^h)\mathbb{E}_{\bar{\theta}^{h+1}}[u(x^{h+1},a_m;\bar{\theta}^{h+1})] \\
=&\sum_{s^{hm+1},\cdots,s^{hm+m-1}\in\mathscr{S}} \Pr(s^{hm+1},\cdots,s^{hm+m-1}|s^{hm};\bar{\theta}^h) \\
\cdot &\sum_{s^{(h+1)m}\in\mathscr{S}} \mathrm{Tr}(s^{(h+1)m}|s^{hm+m-1};\theta^{hm+m-1})\cdot\mathbb{E}_{\bar{\theta}^{h+1}}[\tilde{u}(s^{(h+1)m},a_m;\bar{\theta}^{h+1})].
\end{aligned}$$

Based on the Markov property, we have

$$\begin{aligned}
&\sum_{s^{hm+1},\cdots,s^{hm+m-1}\in\mathscr{S}} \Pr(s^{hm+1},\cdots,s^{hm+m-1}|s^{hm};\bar{\theta}^h)\mathrm{Tr}(s^{(h+1)m}|s^{hm+m-1};\theta^{hm+m-1}) \\
&\qquad = \Pr(s^{(h+1)m}|s^{hm};\bar{\theta}^h).
\end{aligned}$$

Therefore, we obtain (11) by plugging (7) into the definition of ACC in (9). We obtain (12) by taking expectation over $\bar{\theta}^h$ and using the definition of EACC in (10). Based on (13), we can compute $\mathbb{E}_{\bar{\theta}^h}[\Pr(s^{(h+1)m}|s^{hm};\bar{\theta}^h)]$ directly from the transition probability Tr by the forward Kolmogorov equation. □

Remark 1 ***(Computational Equivalency).*** To compute $\tilde{u}$, we generally need to first compute u via (7) and then take expectation over $s^{hm+1},\cdots,s^{hm+m-1}$. This computation is of high temporal and spatial complexity as u depends on x^h. However, for the special case where z is independent of s^k and θ^k, we can compute $\tilde{u}$ directly based on (11) and reduce the computational complexity. Thus, Theorem 1 establishes a computational equivalency between the two DP representations in (7) and (11), which contributes to a lightweight computation scheme. Analogously, we also establish a computational equivalency between the two DP representations in (8) and (12) by taking expectations of (7) and (11) with respect to $\bar{\theta}^h$.

3.4 Data-Driven Assessment

In practice, we do not know the parameters of the SMP model, including the transition probability Tr, the PDF z, the threshold set $\mathscr{N}(s^k,\theta^k)$, and the set of probability of making correct (resp. incorrect) decisions $\bar{p}^n_{CD}$ (resp. $\bar{p}^n_{ID}$), $n\in\{0,1\cdots,N\}$. Therefore, we use Temporal-Difference (TD) learning [15] to evaluate the performance of the AM strategy $a_m\in\mathscr{A}$ based on the inspection results in real-time.

Consolidated IDoS Risk Learning. Letting $v^h(x^h,a_m)$ be the estimated value of $\hat{u}(x^h,a_m)$ at the inspection stage $h\in\mathbb{Z}^{0+}$, we have the following recursive update in real-time as shown in (14).

$$v^{h+1}(\hat{x}^h,a_m)=(1-\alpha^h(\hat{x}^h))v^h(\hat{x}^h,a_m)+\alpha^h(\hat{x}^h)(\hat{c}^h+\gamma v^h(\hat{x}^{h+1},a_m)), \tag{14}$$

where $\hat{x}^h$ (resp. $\hat{x}^{h+1}$) is the observed state value at the current inspection stage h (resp. the next inspection stage $h+1$), $\alpha^h(\hat{x}^h)\in(0,1)$ is the learning rate, and $\hat{c}^h$ is the

observed cost at stage $h \in \mathbb{Z}^{0+}$. To guarantee that v^{∞} convergences to $\hat{u}$, we require $\sum_{h=0}^{\infty} \alpha^h(x^h) = \infty$ and $\sum_{h=0}^{\infty} (\alpha^h(x^h))^2 < \infty$ for all $x^h \in \mathscr{X}$.

Aggregated IDoS Risk Learning. For the special case where PDF z is independent of s^k and θ^k, we can use TD learning to directly estimate EACC $\bar{u}(s^{hm}, a_m)$ based on (12). Letting $\bar{v}^h(x^h, a_m)$ be the estimated value of $\bar{u}(s^{hm}, a_m)$ at the inspection stage $h \in \mathbb{Z}^{0+}$, we have the following recursive update in real-time as shown in (15).

$$\bar{v}^{h+1}(\hat{s}^{hm}, a_m) = (1 - \bar{\alpha}^h(\hat{s}^{hm}))v^h(\hat{s}^{hm}, a_m) + \bar{\alpha}^h(\hat{s}^{hm})(\hat{c}^h + \gamma \bar{v}^h(\hat{s}^{(h+1)m}, a_m)), \quad (15)$$

where $\hat{s}^{hm}$ (resp. $\hat{s}^{(h+1)m}$) is the observed state value at the current inspection stage h (resp. the next inspection stage $h+1$), $\bar{\alpha}^h(\hat{s}^{hm}) \in (0,1)$ is the learning rate, and $\hat{c}^h$ is the observed cost at stage $h \in \mathbb{Z}^{0+}$. To guarantee that $\bar{v}^{\infty}$ convergences to $\bar{u}$, we require $\sum_{h=0}^{\infty} \bar{\alpha}^h(s^{hm}) = \infty$ and $\sum_{h=0}^{\infty} (\bar{\alpha}^h(s^{hm}))^2 < \infty$ for all $s^{hm} \in \mathscr{S}$.

4 Numerical Experiments and Analysis

We provide a numerical case study in this section to corroborate the results. Let the set of category label $\mathscr{S} = \{s_{AL}, s_{NL}, s_{PL}\}$ be the location of the attacks where s_{AL}, s_{NL}, and s_{PL} represent the application layer, network layer, and physical layer, respectively. We consider the special case where $\tau^k, \forall k \in \mathbb{Z}^{0+}$, is an exponential random variable with a constant rate $\beta > 0$, i.e., $z(\tau | s^k, \theta^k) = \beta e^{-\beta \tau}, \forall s^k \in \mathscr{S}, \theta^k \in \Theta, \tau \in [0, \infty)$. Figure 7 illustrates an exemplary sequential attack where the vertical dashed lines represent the attack stages $k \in \mathbb{Z}^{0+}$. The length of the rectangles between the k-th and $(k+1)$-th vertical dash lines represents the k-th attack's duration τ^k. The height of each square distinguishes the attack's type; i.e., tall and short rectangles represent feints and real attacks, respectively.

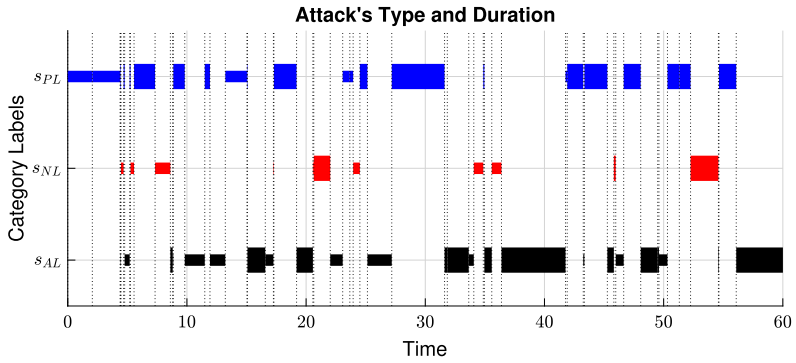


Fig. 7. The sequential arrival of feints and real attacks with different category labels.

The inspection time $\tau_{IN}^{h,m}$, as the summation of m i.i.d. exponential random variables, is an Erlang distribution with shape m and and rate $\beta > 0$, i.e., $\bar{z}(\tau|x^h;\bar{\theta}^h) = \frac{\beta^m \tau^{m-1} e^{-\beta\tau}}{(m-1)!}, \forall x^h \in \mathscr{X}, \forall \bar{\theta}^h, \tau \in [0,\infty)$. Consider a single threshold $N = 1$ and $\mathscr{N} = \{\bar{\tau}_0(s^k,\theta^k), \bar{\tau}_N(s^k,\theta^k)\}$. Then, $\Pr(w^{hm}|x^h,a_m;\bar{\theta}^h)$ in (4) has the following closed form in (16) for correct decisions, i.e., $\theta^{hm} = \theta_{FE}, w^{hm} = w_{FE}$ or $\theta^{hm} = \theta_{RE}, w^{hm} = w_{RE}$.

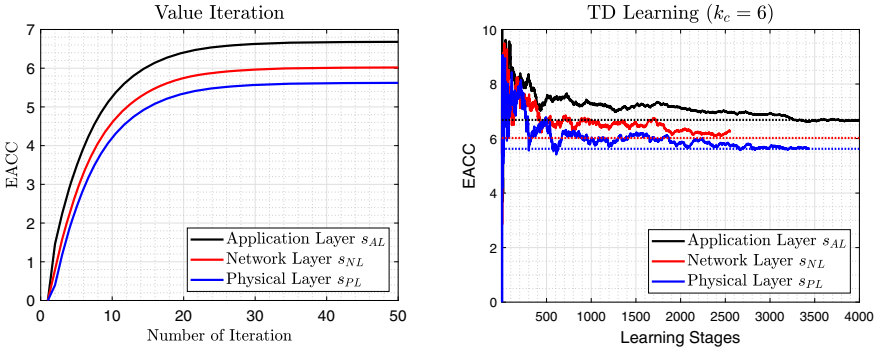
$$\begin{aligned}\Pr(w^{hm}|s^{hm},a_m;\theta^{hm}) &= \int_{\bar{\tau}_N(s^{hm},\theta^{hm})}^{\infty} \bar{p}_{CD}^N(s^{hm},\theta^{hm}) \frac{\beta^m \tau^{m-1} e^{-\beta\tau}}{(m-1)!} d\tau \\ &= \bar{p}_{CD}^N(s^{hm},\theta^{hm})(1 - CDF(\bar{\tau}_N(s^{hm},\theta^{hm}))),\end{aligned} \tag{16}$$

where the Cumulative Distribution Function (CDF) of the random variable $\tau_{IN}^{h,m}$ is

$$CDF(\bar{\tau}_N(s^{hm},\theta^{hm})) = 1 - \sum_{n=0}^{m-1} \frac{1}{n!} e^{-\beta\bar{\tau}_N(s^{hm},\theta^{hm})} (\beta\bar{\tau}_N(s^{hm},\theta^{hm}))^n. \tag{17}$$

4.1 Value Iteration and TD Learning

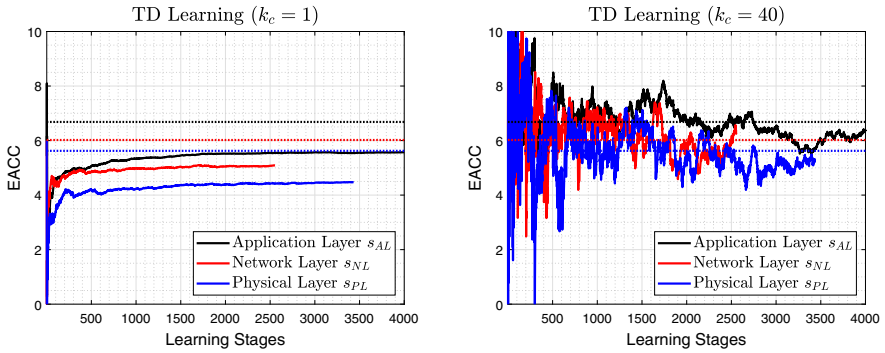
Since PDF z is independent of s^k and θ^k, we can compute EACC in (12) by value iteration. As shown in Fig. 8a, the estimated values of EACC under three different category labels, i.e., $\bar{u}(s_{AL},a_m)$, $\bar{u}(s_{NL},a_m)$, and $\bar{u}(s_{PL},a_m)$ in black, red, and blue, respectively, all converge within 40 iterations. When the exact model is unknown, we use TD learning in (15) to estimate EACC $\bar{u}(s^{hm,a_m})$. In particular, we choose $\bar{\alpha}^h(s^{hm}) = \frac{k_c}{k_{TI}(s^{hm})-1+k_c}$ as the learning rate where $k_c \in (0,\infty)$ is a constant parameter and $k_{TI}(s^{hm}) \in \mathbb{Z}^{0+}$ is the number of visits to $s^{hm} \in \mathscr{S}$ up to stage $h \in \mathbb{Z}^{0+}$. We illustrate the convergence of TD learning in Fig. 8b with $k_c = 6$. Since the number of visits to s_{AL}, s_{NL}, and s_{PL} depends on the transition probability $\bar{T}r$, the learning stages for three category labels are of different lengths.



(a) Theoretical value by value iteration.

(b) Simulated value by TD learning.

Fig. 8. Computation and learning of EACC.



(a) Learning rate decreases too fast $k_c = 1$. (b) Learning rate decreases too slow $k_c = 40$.

Fig. 9. Improper values of k_c lead to unsatisfactory learning performances in finite steps.

If k_c is too small as shown in Fig. 9a, the learning rate decreases so fast that new observed samples hardly update the estimated value. Then, it takes longer learning stages to learn the correct value. On the contrary, if k_c is too large as shown in Fig. 9b, the learning rate decreases so slow that new samples contribute significantly to the current estimated value, which causes a large variation and a slow convergence.

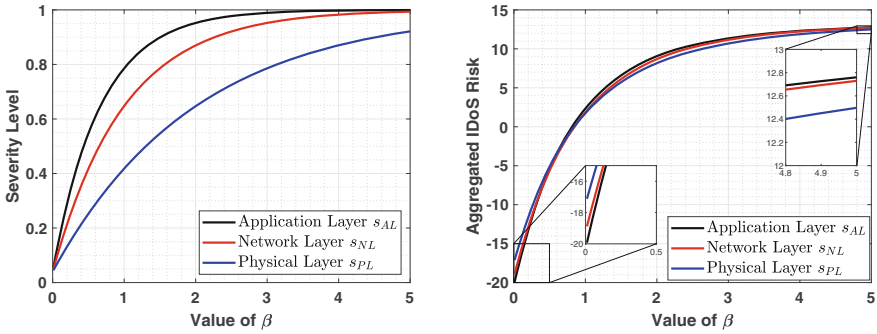
4.2 Severity Level and Aggregated Risk Without Attention Management

When there are no AM strategies, i.e., $m = 1$, the human operator switches attention whenever a new attack arrives. Then, (17) can be simplified as $CDF(\bar{\tau}_N(s^{hm}, \theta^{hm})) = 1 - e^{-\beta \bar{\tau}_N(s^{hm}, \theta^{hm})}, \forall s^{hm} \in \mathscr{S}, \theta^{hm} \in \Theta$, which is an exponential function of the product of the rate $\beta > 0$ and the threshold $\bar{\tau}_N(s^{hm}, \theta^{hm}) > 0$. Thus, $\hat{p}_{CD}(x^h, a_m)$ in (5) decreases monotonously as the value of the product $\beta \bar{\tau}_N(s^{hm}, \theta^{hm})$ increases. Based on Lemma 2, we can write the consolidated severity level as $1 - \hat{p}_{CD}(s^{hm}, a_m)$ without loss of generality. Let $\bar{\tau}_N(s_{AL}, \theta^{hm}) \geq \bar{\tau}_N(s_{NL}, \theta^{hm}) \geq \bar{\tau}_N(s_{PL}, \theta^{hm})$, we plot the severity level, i.e., $1 - \hat{p}_{CD}(s^{hm}, a_m)$, for different values of rate $\beta \in (0,5)$ in Fig. 10a. We illustrate the aggregated IDoS risk versus $\beta \in (0,5)$ in Fig. 10b. As magnified by two insert boxes, the aggregated IDoS risk under s_{AL}, s_{NL}, and s_{PL} can change orders for different β.

4.3 Severity Level and Aggregated Risk with Attention Management

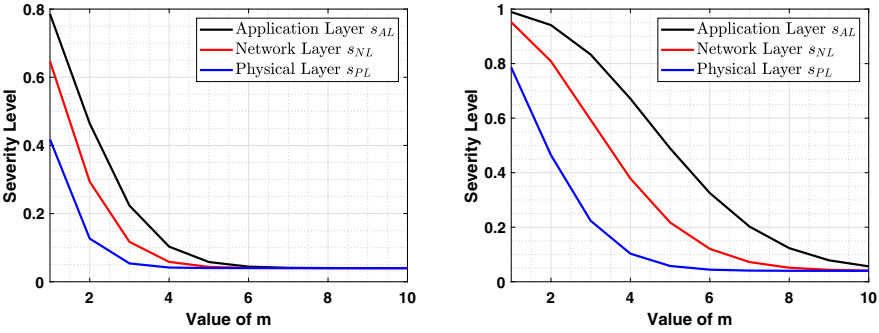
We illustrate how different AM strategies affect the severity level and the aggregated risk of IDoS attacks in Fig. 11 and Fig. 12, respectively, where $\bar{p}^N_{CD}(s^{hm}, \theta_{FE}) = 1$ and $\bar{p}^N_{CD}(s^{hm}, \theta_{RE}) = 0.9$ for all $s^{hm} \in \mathscr{S}$, and $b_{FE} = 0.6$. As shown in Fig. 11, the severity level strictly decreases to 0.04 as m increases regardless of different values of β. We choose a small arrival rate $\beta = 1$ in Fig. 11a and a large rate $\beta = 3$ in Fig. 11b. For a given $m \in \mathbb{Z}^+$, a larger arrival rate results in a higher severity level, and more alerts need to be made inconspicuous to reduce the severity level.

We choose $\beta = 1$ and observe the linear increase of the aggregated IDoS risk when m is sufficiently large in Fig. 11. We investigate how high and low uncertainty costs



(a) Severity level vs. β. (b) Aggregated risk vs. β.

Fig. 10. Severity level and aggregated risk of IDoS attacks under s_{AL}, s_{NL}, and s_{PL} in black, red, and blue, respectively. The insert boxes magnify the selected areas. (Color figure online)

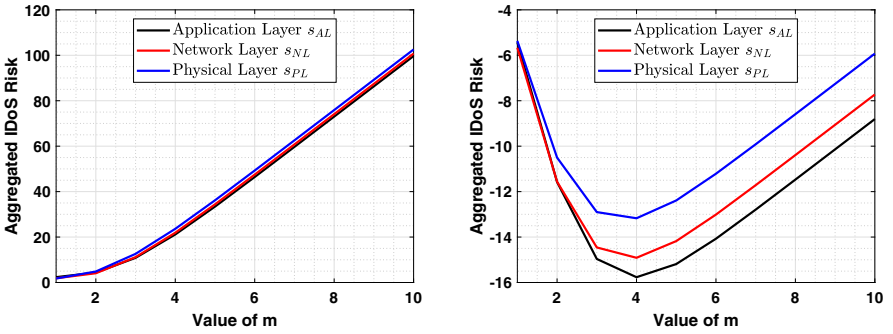


(a) Small arrival rate $\beta = 1$. (b) Large arrival rate $\beta = 3$.

Fig. 11. Severity levels of IDoS attacks under s_{AL}, s_{NL}, and s_{PL} in black, red, and blue. (Color figure online)

c_{UN} affect the aggregated IDoS risk in Fig. 12a and Fig. 12b, respectively. If the uncertainty cost is much higher than the expected reward of correct decision-making, then the detailed inspection and correct security decisions are not of priority. As a result, the 'spray and pray' strategy should be adopted; i.e., let the operator inspect as many alerts as possible and use the high quantity to compensate for the low quality of these inspections. Under this scenario, $\bar{u}(s^{hm}, a_m)$ increases with $m \in \mathbb{Z}^+$ for all $s^{hm} \in \mathscr{S}$ as shown in Fig. 12a. If the uncertainty cost is of the same order as the inspection reward on average, then increasing m in a certain range (e.g., $m \in \{1,2,3,4,\}$ in Fig. 12b) can increase the probability of correct decision-making and reduce the aggregated IDoS risk. The loss of alert omissions outweighs the gain of detailed inspection when m is beyond that range.

Remark 2 **(Rational Risk-Reduction Inattention).** In Fig. 12b, a small m represents a coarse inspection with a large number of alerts while a large m represents a fine



(a) High cost $c_{UN}(s^{hm}, \theta^k) = 20, \forall s^{hm}, \theta^{hm}$. (b) Low cost $c_{UN}(s^{hm}, \theta^k) = 0.2, \forall s^{hm}, \theta^{hm}$.

Fig. 12. Aggregated IDoS risks under s_{AL}, s_{NL}, and s_{PL} in black, red, and blue. (Color figure online)

inspection of a small number of alerts. The U-shape curve reflects that the minimum risk is achieved with a proper level of *intentional inattention* to alerts, which we refer to as the *law of rational risk-reduction inattention*.

5 Conclusion

Attentional human vulnerability can be exploited by attackers and leads to a new class of advanced attacks called the Informational Denial-of-Service (IDoS) attacks. IDoS attacks intensify the shortage of human operators' cognitive resources in this age of information explosion by generating a large number of feint attacks. These feints distract operators from detailed inspections of the alerts, which significantly decrease the accuracy of their security decisions and undermine cybersecurity. We have formally introduced the IDoS attacks and established a quantitative framework that provides a theoretic underpinning to the IDoS attacks under limited attention resources. We have developed human-assistive security technologies that intentionally make selected alerts inconspicuous so that human operators can pay sustained attention to critical alerts.

We have modeled the sequential arrival of IDoS attacks as a semi-Markov process and the probability of correct decision-making as an increasing step function concerning the inspection time. Dynamic Programming (DP) and Temporal-Difference (TD) learning have been used to represent long-term costs and evaluate human performance in real-time, respectively. We have established the *computational equivalency* between the DP representation of the Cumulative Cost (CC) (resp. Expected Cumulative Cost (ECC)) and the Aggregated Cumulative Cost (ACC) (resp. Expected Aggregated Cumulative Cost (EACC)). This equivalency has reduced the dimension of the state space and the computational complexity of the value iteration and online learning algorithms.

From the case study, we have validated that both the severity level and the aggregated risk of IDoS attacks increase exponentially with the product of the attack's arrival

rate and the operator's inspection efficiency. When Attention Management (AM) strategies are applied, we have observed that the severity level strictly decreases with the inspection time. We have arrived at the 'less is more' security principle in cases where correctly identifying the real and feint attacks is of high priority. It has been shown that inspecting a small number of selected alerts with sustained attention outperforms dividing the limited attention to inspect all alerts.

The future work would focus on coordinating multiple human operators to share the cognition load. Based on the literature of cognitive science and existing results of human experiments, we would develop detailed models of human attention, reasoning, and risk-perceiving to better characterize human factors in cybersecurity. Finally, we would extend the periodic AM strategies to adaptive ones that use the feedback of the alerts' category labels and the operator's current cognition status reflected by biosensors.

References

1. Hitzel, B.: The art of cyber war and cyber battle: deception operations (2019). https://www.networkdefenseblog.com/post/art-of-cyber-war-deception
2. Wickens, C.D., Hollands, J.G., Banbury, S., Parasuraman, R.: Engineering Psychology and Human Performance. Psychology Press, London (2015)
3. Casey, W., Morales, J.A., Wright, E., Zhu, Q., Mishra, B.: Compliance signaling games: toward modeling the deterrence of insider threats. Comput. Math. Organ. Theory **22**(3), 318–349 (2016). https://doi.org/10.1007/s10588-016-9221-5
4. Huang, L., Zhu, Q.: Duplicity games for deception design with an application to insider threat mitigation. IEEE Trans. Inf. Forensics Secur., arXiv preprint arXiv:2006.07942 (2021). https://doi.org/10.1109/TIFS.2021.3118886
5. Huang, L., Zhu, Q.: INADVERT: an interactive and adaptive counterdeception platform for attention enhancement and phishing prevention, arXiv preprint arXiv:2106.06907 (2021)
6. Huang, Y., Huang, L., Zhu, Q.: Reinforcement learning for feedback-enabled cyber resilience, arXiv preprint arXiv:2107.00783 (2021)
7. Huang, L., Zhu, Q.: A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems. Comput. Secur. **89**, 101660 (2020)
8. Huang, L., Zhu, Q.: A dynamic game framework for rational and persistent robot deception with an application to deceptive pursuit-evasion. IEEE Trans. Autom. Sci. Eng., 1–15 (2021). https://doi.org/10.1109/TASE.2021.3097286
9. Huang, L., Zhu, Q.: Adaptive honeypot engagement through reinforcement learning of semi-Markov decision processes. In: Alpcan, T., Vorobeychik, Y., Baras, J.S., Dán, G. (eds.) GameSec 2019. LNCS, vol. 11836, pp. 196–216. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32430-8_13
10. Zhao, D., et al.: Bidirectional RNN-based few-shot training for detecting multi-stage attack, arXiv preprint arXiv:1905.03454 (2019)
11. Sims, C.A.: Implications of rational inattention. J. Monet. Econ. **50**(3), 665–690 (2003)
12. Kahneman, D.: Thinking, Fast and Slow. Macmillan, London (2011)
13. Schwartz, B.: The Paradox of Choice: Why More is Less. Ecco, New York (2004)
14. Ban, T., Samuel, N., Takahashi, T., Inoue, D.: Combat security alert fatigue with AI-assisted techniques. In: Cyber Security Experimentation and Test Workshop, pp. 9–16 (2021)
15. Bertsekas, D., Tsitsiklis, J.: Neuro-Dynamic Programming, vol. 27 (1996)