Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

A TESTING BASED APPROACH FOR SECURITY ANALYSIS OF SMART SEMICONDUCTOR SYSTEMS

Robert Dodge
Giulia Pedrielli

School of Computing and Augmented Intelligence
Arizona State University
699 S Mill Ave
Tempe, AZ 85281, USA

Petar Jevtić

School of Mathematical and Statistical Sciences
Arizona State University
900 Palm Walk
Tempe, AZ 85281, USA

2023 Winter Simulation Conference (WSC) | 979-8-3503-6966-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/WSC60868.2023.10408730

ABSTRACT

Digital factories have been recognized as a paradigm with considerable promise for improving manufacturing performance. Digital Twins have emerged as a powerful tool to improve control performance for large-scale smart manufacturing systems. We argue that DT-based smart factories are vulnerable to attacks that use the DT to damage the system while remaining undetectable, specifically in high-cost processes, where DT technologies are more likely to be deployed. As an instructive example, we look into smart semiconductor processes with focus on photolithography. To this end, we formulate a static optimization problem to maximize the damage of a cyber-attack against a photolithography digital twin that minimizes detectability to the process controller. Results demonstrate that this problem formulation provides attack policies that successfully reduce the throughput of the system at trade off of increased detectability to a common process control technique. Results encourage more research in the domain, especially to face scalability and policy-like solutions.

1 INTRODUCTION

Digital factories have been recognized as a paradigm with a considerable promise for improving manufacturing performance, and both industry and public funding have been dedicated to methodologies and tools/technologies to support its development. Industrial Internet of Things (IIoT) and artificial intelligence have been increasingly adopted on the factory floor, leading to *smart factories* with connected machines and operators, and simulations in-the-loop. As an example, sensors, with their associated connectivity and the cloud storage to maintain data, have led to a revolutionary impact on performance improvement, preventive maintenance, and asset management (Technology and Services Industry Association 2022; Link Labs IoT Technology 2023; NOKIA Solutions 2023). It is less clear to what extent these technologies can improve *system-level* performance, particularly in large-scale factories, such as in semiconductors with its thousands of manufacturing processes and sub-processes and what are potential risks. In fact, DTs (Technology and Services Industry Association 2022; Jiang et al. 2021; Schleich et al. 2017; Negri et al. 2017; Li et al. 2022) have emerged as a potential solution to embed and integrate sensor data into factory decision-making by providing model-driven augmentations of these data through simulation, optimization, and forecasting.

We argue that a fundamental aspect to consider for DT-based smart factories implementations is that the system is made vulnerable to attacks that can use the DT information and be hard to detect while damaging the system performance.

The United States Department of Homeland security investigated 63 cyber-attacks against critical manufacturers in 2016 (United States Department of Homeland Security 2016). In 2017, Renault-Nissan was one of numerous organizations that suffered as part of the WannaCry ransomware attack. In this attack,

five manufacturing facilities were forced into a shutdown for several days (Eisenstein, Paul 2017). IBM estimates that the average cost of a data breach in 2022 at $4.35 million, a 13% increase from 2020 (IBM 2022). Beyond data breaches, damage to work-in-process items or equipment is possible in some extreme cases. During the Stuxnet attack in 2012 a computer worm was utilized to damage nuclear centrifuges by targeting their control systems (Sturm et al. 2014). This attack is estimated to have ruined approximately one-fifth of Iran's nuclear centrifuges (Kelley, Michael 2013). The risks presented by these attacks is unacceptable for many high-end manufacturing facilities. We argue that this is particularly the case for facilities that have high value assets that fabricate critical products.

This paper will focus on the potential risks that cyber-attacks pose to cyber-physical digital twin controlled semiconductor photolithography processes, when the cyber-attack specifically concerns the digital twin. Due to the high costs of these processes and high level of automation present in photolithography (Byrne 2007), these vulnerabilities have come under increased scrutiny due to upticks in cyber-attacks. In particular, we focus on *stealthy* attacks, i.e., situations where the anomaly detection cannot identify the attack. To achieve this, we will utilize the DTFab simulation model implemented in (Pedrielli et al. 2023) to examine the possibility for a cyber-attacker to stealthily attack a semiconductor wafer photolithography process. This model is a discrete event simulation coded in Python (version 3.88) using the SimPy library. This model serves as a simulated replica of a physical photolitography process that mimics the behavior of true system. In this model, lots composed of homogeneous wafers move through a photolithography process completing a series of operations (develop, bake, scan and several manipulations) before departing the system. To complete the scan operation, which is responsible to impress the wafers, a number of identical photolitography machines (each of a value ranging from 100 to 200 Million dollars) operate using mounted reticles in parallel and perform operations. When a lot has completed processing, the digital twin performs a reticle switch between the idle tool and the main storage before the next lot begins processing.

The primary purpose of the DTFab twin is to handle the allocation of reticles and product lots to stations. Hence, we will assume that an attacker is able to influence these functions of the physical system. In particular, we will assume that the access prevents the attacker from doing physical damage to the equipment or product, however, it will be able to "fake" states that are relevant to the reticle management system of the photolithography process. Hence, the damage is largely attributable to the capability of the attacker to generate inefficiencies in the flow of wafers and the flow or reticles in a way that damages the system throughput (Benzoni et al. 2020; Byrne 2007). Specifically, we will assume that the attacker's goal is to hinder the process long-term throughput while minimizing detectability. To measure this detectability, we will assume that the system utilizes a process control chart operated on the physical layer of the system. We utilize this method because prior literature has already discussed the risks of a compromised digital twin that is used for anomaly detection (Danilczyk et al. 2021; Sahal et al. 2021). The exact implementation details for this model can be found in (Pedrielli et al. 2023).

The approach taken to examine this problem as well as justifications for various decisions can be found in Section 3.

2 LITERATURE REVIEW

A rich literature is dedicated to the study of cyber-attacks with a variety of approaches due to the wide variety of attacks and defense mechanisms that have been proposed and currently are employed in industry. Because of this, we organized the literature analysis intro three focus areas: (i) First we review approaches for cyber-attacks on general IoT devices, then (ii) we review methodologies for attacks on cyber-physical systems in particular, and finally (iii) we discuss the specific dangers that these attacks pose on manufacturing systems utilizing digital twin technology.

2.1 General IoT Cyber-attacks

Many industries such as healthcare, industrial manufacturing, and vehicle automation utilize IoT systems (Misra and Saha 2019). Since a compromise of the cyber-security layer of the system can have disastrous consequences, a great quantity of literature is dedicated to examining methods that can protect against these attacks. Alsamiri and Alsubhi (2019) discuss the use of machine learning algorithms to detect cyber-attacks within IoT systems. The authors specifically note that these algorithms can be divided into two detection methods, anomaly based and signature based. Anomaly based detection methods observe network traffic and mark suspicious traffic as an attack. They observe that a problem with this type of machine learning detection is the potential of a false positive. Uncommon traffic patterns for a system, which may be entirely legal, may result in the machine learning agent tagging the traffic as an attack. Examples of this approach can be seen in the works by Baig et al. (2020), and Koroniotis, Moustafa, Sitnikova, and Turnbull (2019). While these machine learning algorithms are able to change their behavior and learn with new data, they are highly resource intensive (Garcia-Teodoro et al. 2009). The signature based approach is discussed in the works by Hubballi and Suryanarayanan (2014) and Ioulianou et al. (2018). These signature based approaches are successful at detecting known attacks, but struggle to identify new kinds of attacks.

2.2 Cyber-attacks Against Cyber-Physical Systems

An example of this type of attack was the Stuxnet attack against the centrifuges Iran used to refine uranium (Sturm et al. 2014). Where, the attacker was able to gain control of the centrifuges through the controlling software and feed falsified data back to the controller. Damaging the equipment, while making it seem as if the equipment was all clear from the perspective of the operator (Kelley, Michael 2013).

Additive manufacturing is one sector that has dealt with the potential of cyber-attacks such as these. The risks of these attacks such as damaged final product, damaged equipment, or intellectual property theft, are discussed in Yampolskiy et al. (2015). Turner et al. (2015) presents several different attack approaches and tests which defense measures and attacks were most efficacious. Zeltmann et al. (2016) examines the ability to hide embedded defects in a compromised additive manufacturing process. While the defect was undetectable to an ultrasonic inspection, the defect did not produce a significant decrease in tensile strength.

2.3 Cyber-attacks Against Digital Twin Systems

Prior research about digital twins and attack detection has been conducted in the context of utilizing the digital twin as an anomaly detection method to detect when an attack has occurred. Xu et al. (2021) implements a machine learning algorithm to work with a continuously built digital twin called ATTAIN. Zhao et al. (2022) presents a framework where the state of the physical system and digital twin simulation are continually synchronized and differences between the simulated state and physical state are observed. These differences can be observed and analyzed to attempt to detect anomalies. Eckhart and Ekelhart (2018) demonstrate the ability for a digital twin to be used as an experimental platform for attacks on a CPS. Beyond attempting to detect malicious attacks, digital twins are also used to detect general process anomalies from deteriorating equipment or a system shifts. Sahal et al. (2021) uses this method to detect erratic behavior in a network of wind turbines. Danilczyk et al. (2021), uses digital twin technology to detect anomalies in a sensor network. These approaches highlight a serious security issue. An attacker who gains control of the digital layer of the system, also gains control of the anomaly detection, and is thus able to circumvent the physical system's defenses.

Research on attacks against the digital twin component of the system is less extensive. Eckhart and Ekelhart (2019) additionally discuss the use of a compromised digital twin to control the physical elements of the system. In a compromised digital twin, an attacker could hide their attack on the physical system through an attack on the digital half. Another common use case for digital twins is as a comparison

point for process data. In this situation, a compromised digital twin could result in a Garbage In Garbage Out situation, where the perceived baseline data created by the digital twin is bad from the start. Suhail et al. (2022) discuss the potential for a blockchain based digital twin, where a blockchain's decentralized nature makes it resilient to unauthorized alterations. They note that this does not entirely prevent the attack scenario, a bad actor having access to the system could result in bad data being recorded to the blockchain in the first place.

2.4 Contribution

Preexisting literature has clearly demonstrated the risks associated with a compromised digital twin. These risks pose serious questions for industries seeking to enhance their production systems through the use of this technology. In this paper, we focus on the semiconductor manufacturing industry. The current semiconductor industry exists at a specific intersection of possessing both highly expensive equipment that is also highly automated. As the semiconductor industry frequently employs IoT to manage its hardware, and has recently made advancements to incorporate digital twins, there is substantive risk should the security of the system be compromised. Because of this, we wish to explore the potential for an attacker who has gained access to the controlling digital twin to hinder the throughput of the process while remaining undetected by a process control chart.

3 PROPOSED APPROACH

As we are attempting to analyze the vulnerabilities of a system with a compromised digital twin, we will assume that an attacker has gained access to the digital layer of system *a priori*. We will first discuss the problem from a general intuitive sense, then present a mathematical formulation and a proposed solution method. In our formulation, an attacker who has gained (partial) access to the digital twin is able to access the state (physical and simulated) of the tools and reticles (masks) in the system. With this in mind, the attacker goal is to systematically reduce the throughput of the system feeding false information about the state of the equipment and reticles through the DTFab digital twin. The remainder of this section will be dedicated to formulating the problem as a static optimization, providing justification for choices made in formulation, and the selected solution methodology.

3.1 Notation

In DTFab, the state of the system is represented as the set of reticle, lot, and wafer objects that also contain all the relevant unique identifying information regarding the system. Formally, the state of the system at step k, referred to as S_k, is a set such that, $S_k = \{P_k, L_k, E_k, W_k\}$, where P_k, L_k, E_k, and W_k refer to sets of reticles, lots, tools, and wafers, respectively, and they are indexed by $a = 1,\ldots,N^p$, $b = 1,\ldots,N^l$, $c = 1,\ldots,N^e$, and $d = 1,\ldots,N^w$, respectively representing the number of reticles, lots, tools, and wafers. Thus, we have the following: $P_k = \{p_{a,k}\}, a = 1,\ldots,N^p; L_k = \{l_{b,k}\}, b = 1,\ldots,N^l; E_k = \{e_{c,k}\}, c = 1,\ldots,N^e; W_k = \{w_{d,k}\}, d = 1,\ldots,N^w$.

In the attack problem formulation we will consider T as the observation horizon (this will be the run length of the simulations in the experiment setup), while we will refer to τ as the duration of an individual attack from a nefarious agent. Since our experiments are synthetic in nature, we will always consider a ν interval of time referred to as *warm up* during which no attack can be implemented. This will work as a needed time to obtain steady state representative system observations.

3.2 Threat Model

In this section we introduce the threat model that we use in the attempt to inflict *system-level stealthy attacks* against the photolithography system. Our threat model has two key components: (i) we first define

the relevant variables to describe the attacker access to the system; (ii) we then define the reward function used to direct the search of the optimal attack decision.

Attacker Model. We will consider a system with a number of tools, N^e, and a number of reticles, N^p. During the process, lots enter the system for processing, each requiring N^o operations before departure. Our objective is to identify an optimal attack over a finite time horizon of length T. Once the warm up of length ν is completed, the attacks will begin, lasting until the end of the horizon. While control-based formulations of this problem are possible, in this first exploration we will propose a one shot optimization as the attack tactic of the malicious actor. In order to do so, we assume that the attacker can decide to perform an attack at a frequency $1/\tau$ to be optimized as part of the procedure. Equivalently, we will generate a list $U = \{u_k\}$, consisting of $\frac{\mathrm{T}-\nu}{\tau}$ attacks. Each attack u_k, $k = 0,\ldots,\frac{\mathrm{T}-\nu}{\tau}$, will be comprised of a tool attack and/or a reticle attack on any subset of reticles or tools in $P_k \cup E_k$. These attacks will be referred to as $u_k = \{X_k^E, X_k^P\}$, where X_k^E and X_k^P are vectors of decision variables representing tool and reticle attacks, respectively, namely:

$$X_k^E = \left\{x_{c,k}^E\right\} : x_{c,k}^E = \begin{cases} 1, & \text{If tool } c \text{ is attacked to show a full queue during attack } u_k \\ 0, & \text{Otherwise} \end{cases}$$

$$\forall\ k,\ c$$

$$X_k^P = \left\{x_{a,k}^P\right\} : x_{a,k}^P = \begin{cases} 1, & \text{If reticle } a \text{ is attacked to be unavailable during attack } u_k \\ 0, & \text{Otherwise} \end{cases}$$

$$\forall\ k,\ a$$

In other words, this means that a tool attack during a time step k would cause the system to falsely believe that the queue of the attacked tool is full, and no new lots can enter the attacked tool's queue for the duration of the time step. This causes the system to be unable to allocate reticles optimally, producing longer cycle times. Similarly, a reticle attack would prevent any new operation using the attacked reticle from beginning for the duration of the attack. This would mean that if a reticle mounted on a photolithography tool is attacked, any new lot requiring the attacked reticle would need to initiate a reticle switch if a replacement is available from storage, or return to the main loading bay, delaying the system.

In our formulation, X_k^E, X_k^P have N^e and N^p dimensions, respectively. This creates a $(N^e + N^p) * \frac{\mathrm{T}-\nu}{\tau}$ dimensional problem. As the average photolithography process can have thousands of reticles (Park et al. 1999), the reticle attack section of the decision vector can produce a space with thousands of dimensions that would be computationally expensive to explore with brute force approaches. Since this paper focuses on steady state behavior of the system, it is not an unreasonable assumption that an attack on the system aiming to decrease its throughput will also have cyclical behavior. Because of this, rather than formulating the entire attack horizon as independently operating avenues of attack, we will formulate the optimization to select an attack frequency for each reticle and tool. This will allow us to solve for an approximately optimal attack policy $\hat{U}$. Letting Y^E and Y^P represent the period decision variables for tools and reticles, respectively, these are collections of actions over the tools and reticles, i.e., $Y^E = \left\{y_c^E\right\}$, and $Y^P = \left\{y_a^P\right\}$,

where the element are defined as follows.

$$y_c^E = \begin{cases} n, & \text{If tool loader } c \text{ is attacked to show a full queue during attack every } n \text{ attack periods} \\ 0, & \text{Otherwise} \end{cases}$$

$$n = 1, \ldots, \frac{\mathrm{T} - \nu}{\tau},\ c = 0, \ldots, N^e$$

$$y_a^P = \begin{cases} m, & \text{If reticle } a \text{ is attacked to be unavailable every } m \text{ attack periods} \\ 0, & \text{Otherwise} \end{cases}$$

$$m = 1, \ldots, \frac{\mathrm{T} - \nu}{\tau},\ a = 0, \ldots, N^p$$

For example, in terms of the introduced variables, this means that if $y_1^E = 2$, tool 1 would be attacked every other attack period. The same statement applies for any value of c or n and any value of a or m. This definition reduces the space to be $(N^e + N^p)$. Once optimal values of Y^P and Y^E are solved for, they can be translated back into an approximately optimal policy $\hat{U}$.

Attack Reward Design. The static optimization problem that will solve for the approximate policy $\hat{U}$ can be designed to consider several forms of cost/reward. In this contribution, the objective function will be to minimize the maximum of the probability that the performance of the system is lowered by a threshold δ and the probability that the attack will be detected by the process control algorithm. These probabilities will be referred to as $\phi_1(U)$ and $\phi_2(U)$, respectively. As both of these measures are random variables, the problem can be formulated as the following optimization:

$$\begin{aligned} &\min_{Y^E, Y^P} \max(\phi_1(U), \phi_2(U)) \\ &\text{s.t. } y_a^P \in \{1, \ldots, \frac{\mathrm{T} - \nu}{\tau}\} \ \forall\ a \\ &\qquad y_c^E \in \{1, \ldots, \frac{\mathrm{T} - \nu}{\tau}\} \ \forall\ c \end{aligned}$$

where $\phi_1(U)$ is the probability that the difference of the expected cycle time of the attacked system and the ideal system is less than a threshold δ, namely:

$$\phi_1(U) = P(E(\theta(U) - \theta^0) < \delta).$$

And ϕ_2 represents the probability that the control chart of the system under attack is outside of its detection threshold.

$$\phi_2(U) = P(\theta(U) \notin R_c)$$

Here, θ represents the cycle time and is inversely proportional to the throughput of the system. R_c represents the control region defined by a normal approximation control chart. Where $R_c = [LCB, UCB]$, such that $LCB = E(\theta) - Z_{\varepsilon/2} \cdot \sigma_\theta$ and $UCB = E(\theta) + Z_{\varepsilon/2} \cdot \sigma_\theta$, where $Z_{\cdot}$ is the quantile of the standard normal distribution. With $E(\theta)$ and σ_θ estimated through long run simulations.

3.3 Solution Approach: Sequential Generation of Attacks

As these probabilities cannot be solved for in closed form, instead they will be evaluated through simulation, thus producing the estimators $\hat{\phi}^1$ and $\hat{\phi}^2$. Since the simulation model has stochastic elements and operates as a black box function, the optimization algorithm chosen to be utilized is an elitist evolutionary algorithm.

Evolutionary Algorithm Logic. At each iteration within this algorithm, a generation of attack policies of size N is generated. For the purposes of notation, the generation for iteration i will be referred to as Ψ_i. Each attack series within Ψ_i is evaluated for efficacy through r replicated simulations. Each of

these policies is evaluated based on the aforementioned, $\hat{\phi}^1$ and $\hat{\phi}^2$. Letting r be the index representing the r-th replication, we have the following

$$\hat{\phi}_1^i(U_i) = \frac{1}{r}\sum_{p=1}^{r} \mathbb{1}(E(\theta(U)_{i,p} - \theta^0) < \delta) \text{ and } \hat{\phi}_2^i(U_i) = \frac{1}{r}\sum_{p=1}^{r} \frac{P_{i,p}}{Z_{i,p}}.$$

Where $P_{i,r}$ and $Z_{i,r}$ represent the number of lots from replication r in iteration i whose cycle time exceeds the bounds defined by R_c and the number of lots who are fully completed before the end of the simulation, respectively. Once each attack from the generation is evaluated, the best attack from that generation is compared to the previous best attack. If the best attack from the new generation has a better objective function result than the previous best, the current best attack and current best objective value are updated accordingly. Then, a selection of the best attacks from the generation are selected and modified for the next generation. These attacks are modified according to a crossover rate and a mutation rate, ζ and ω, respectively. In this implementation, a uniform crossover method is utilized, meaning each element within the decision variable vector is inherited from either parent with equal probability. The mutation rate determines the probability that each value within a solution is replaced by a random feasible value. As this instance of an evolutionary algorithm is an elitist algorithm, an additional proportion of the best of each generation is guaranteed to survive to the next generation according to the proportion parameter γ. These generations will be iteratively generated until either a maximum number of generations has been reached, or a number of generations have passed without improvement, ρ and ξ, respectively. Letting U_i represent the attack series for iteration i, U^* represent the current best attack series, and $f(U^*)$ represent the objective value of the current best attack series. This logic is summarized in Figure 1 and Algorithm 1.

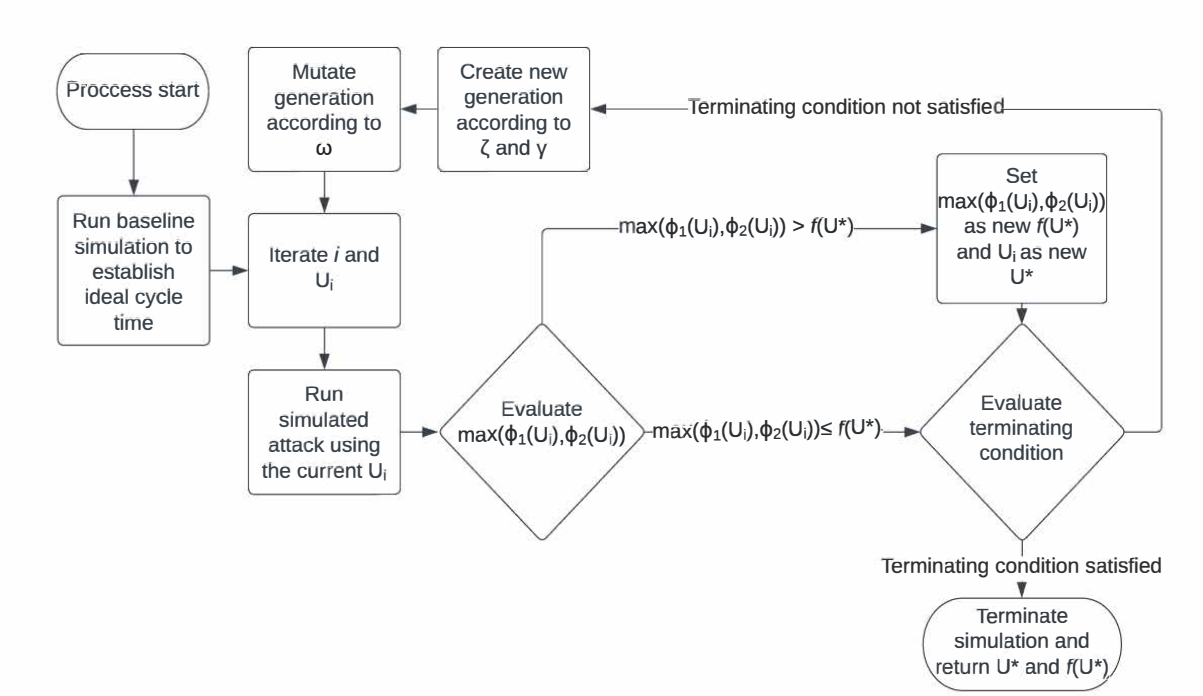

Figure 1: Flow chart illustrating the logic used by the evolutionary algorithm to solve the static optimization problem presented in Section 3.2.

Algorithm 1 Stealthy Attacks for Photolithography with DT in the loop.

1: Run baseline simulation for r replications. Collect the expectation of the cycle time of lots as CT^0.
2: Begin optimization with a boundary condition U^0 and $i = 0$.
3: Iterate i and U_i.
4: Simulate attack on the system using the current policy U_i for r replications.
5: Evaluate objective function for iteration i, $\max(\phi_1(U), \phi_2(U))$.
6: If $\max(\phi_1(U), \phi_2(U)) > f(U^*)$, save U_i and $f(U_i)$ as the new U^* and $f(U^*)$.
7: If terminating condition is satisfied, terminate the optimization algorithm and return U^* and $f(U^*)$ as optimal. Otherwise, continue on to 8.
8: Create the next generation to be evaluated according to crossover rate ζ and elite ratio γ
9: Mutate new generation according to the mutation rate ω. Go to 3

4 RESULTS AND DISCUSSION

Two sets of experiments were performed using the DTFab simulator (Pedrielli et al. 2023). The first set consists of a smaller model with only two photolitography tools for validation and initial exploration. Under this size it was possible to evaluate attacks generated with a grid (*brute force*). The second experimental set features a larger 10-tool model. In this case using brute force is impractical, and the evolutionary approach is used instead.

In the *first set*, two experiments were conducted. For these, a design of experiments was produced. For these experiments, a small scale photolitography model was used. The model operating parameters are shown in Table 1. In these experiments, the effects of the independent variables characterizing the reticle attack frequencies and tool attack frequencies on the detection rate and average cycle time were observed independently. As the system mounts 27 reticles, attacks were grouped by associated product type to reduce the quantity of simulations required. For example, an attack frequency of 16 to reticle P1 implies that all reticles associated with product type 1 were attacked once every 16 time periods.

In the *second set of experiments*, the evolutionary algorithm discussed in Section 3.3 was implemented on a larger system. For the targeted error threshold, a value of 3500 seconds was chosen, this results in an approximately 6% increase in overall cycle time. The parameters for the evolutionary algorithm and system parameters for this round of experiments are shown in Table 2. These parameter values were chosen as a compromise to allow for adequate exploration of the solution space while also allowing the optimization to be run in a practical time frame. As this is a stochastic problem, 15 macroreplications were executed to estimate the noise associate to the estimated attack solution.

Discussion. For the small system experiment, the main effect and interaction plots with respect to the error and anomaly metrics are shown in Figures 2,3, and 4. From the results, we can observe that attacks against the system's reticles quickly cause the percentage of control chart anomalies to rise. Even relatively low attack rates against the reticles for a single product type produces an anomaly rate in excess of 10%. The tool attack experiments however, produces comparable decreases in average cycle time with

Table 1: Parameters for small system experiments.

Parameter	Variable	Value
Number of tools	N^e	2
Warmup period	ν	250000 s
Number of product types	N^m	3
Operations per product type	N^o	3
Copies of each reticle	$\frac{N^p}{N^o * N^m}$	3
Total number of reticles	N^p	27
Attack length	τ	10000 s
Number of attacks	$\frac{T-\nu}{\tau}$	64
Total length of attack horizon	$T-\nu$	640000 s

Table 2: Parameters for large system experiments and evolutionary algorithm.

(a) Large system evolutionary algorithm parameters.

Parameter	Variable	Value
Max number of iterations	ρ	25
Population size	η	75
Mutation probability	ω	0.1
Elite ratio	γ	0.1
Crossover probability	ζ	0.7
Max number of iterations without improvement	ξ	10
Error threshold	δ	3500 s

(b) Large system simulation parameters.

Parameter	Variable	Value
Number of tools	N^e	10
Warmup period	ν	250000 s
Number of product types	N^m	3
Operations per product type	N^o	5
Copies of each reticle	$\frac{N^p}{N^o * N^m}$	3
Total number of reticles	N^p	45
Attack length	τ	10000 s
Number of attacks	$\frac{T-\nu}{\tau}$	100
Total length of attack horizon	$T-\nu$	1000000 s

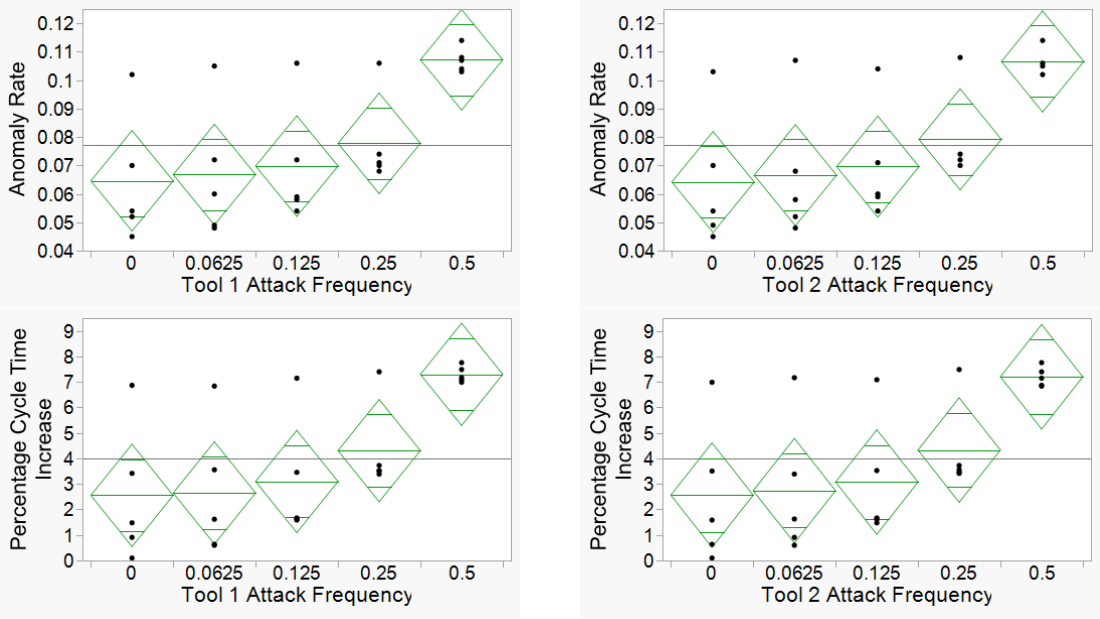

Figure 2: Main effects plots of tool attack frequencies by average cycle time increase and anomaly rate.

a lower chance of being detected. Both attacks frequencies have negative interaction terms. This indicates that attacking multiple reticles/tools simultaneously has diminishing returns with regards to increasing the average cycle time.

For the large model experiments, all macroreplications found a feasible optima that produced the desired 3500 second average slowdown. Each of these optima were unique and each had comparable anomaly rates, with an average anomaly rate of 0.083. These optimization results (see Figure 5) also indicate that, while we still have a rate of detection larger than the desired (5% in this case) the undetectability rate remains above 90% for all the generated attacks. Furthermore, these attacks still successfully increase the average cycle time of the process by more than 5%. From the perspective of an attacker, this may be an entirely desirable result, despite the inevitable eventual discovery of the attack.

5 CONCLUSIONS

In this paper, a digital twin-based smart control for reticle and tool management in a photolithography system was used as an experimentation platform to examine the impacts of a compromised digital twin system. In this regard, we were able to successfully demonstrate the effectiveness of formulating a cyber-attack on

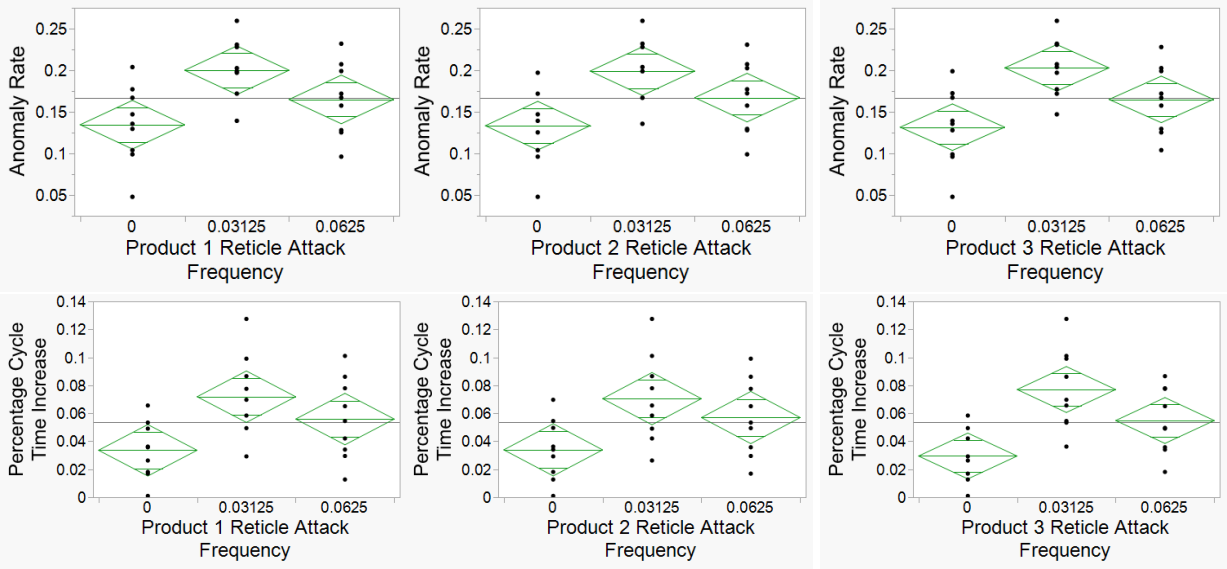

Figure 3: Main effects plots of reticle attack frequencies by average cycle time increase and anomaly rate.

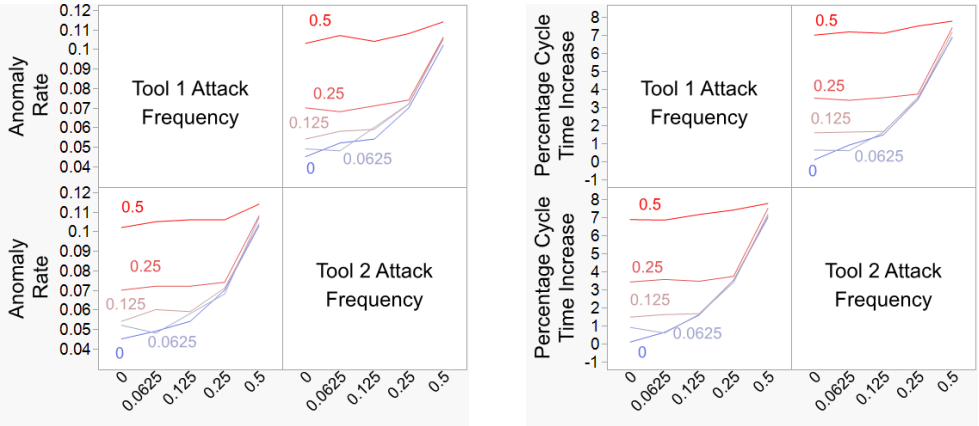

Figure 4: Interaction plot of tool attack frequencies by average cycle time increase and anomaly rate.

said digital twin as a static optimization problem with the goal of hampering the process by a predetermined cycle time increase. We show how an increase of the cycle time still achieve a 90% undetectability rate. Even if this is smaller than the desired 95%, we can argue that even under such conditions, since the control chart will need an extended period of time to detect a low intensity attack, we are capable of slowing the process by several percentage points.

While this paper approaches the security analysis problem from the perspective of a static optimization, we can see the possibility to define this in the context of online or optimal control where attacks are chosen sequentially using information gathered from the previous attack and system responses. Additionally, the large variance in real life photolithography processes may cause scalability issues for the results produced in this paper. It is not uncommon for photolithography manufacturing systems to posses many more reticles than are used in these simulations, meaning an attacker wishing to use this approach would have an exceedingly high dimensional problem that is likely computationally intractable. In this sense approaches for aggregation/parallelization are worthy of exploration.

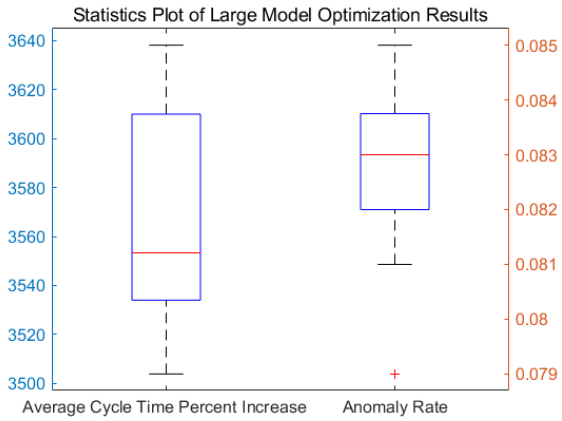

Figure 5: Summary statistics for optimization macroreplication results.

ACKNOWLEDGEMENTS

This work is funded by the National Science Foundation under grant CNS-2000792 and partially supported by the Intel Research grant #00035705.

REFERENCES

Alsamiri, J., and K. Alsubhi. 2019. "Internet of Things Cyber Attacks Detection Using Machine Learning". *International Journal of Advanced Computer Science and Applications* 10(12).

Baig, Z. A., S. Sanguanpong, S. N. Firdous, T. G. Nguyen, C. So-In et al. 2020. "Averaged Dependence Estimators for DoS Attack Detection in IoT Networks". *Future Generation Computer Systems* 102:198–209.

Benzoni, A., C. Yugma, P. Bect, and A. Planchais. 2020. "Allocating Reticles in an Automated Stocker for Semiconductor Manufacturing Facility". In *Proceedings of the 2020 Winter Simulation Conference*, edited by Z. Zheng, K.-H. Bae, S. Lazarova-Molnar, B. Feng, and S. Kim, 1711–1717. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Byrne, P. J. 2007. "An Analysis of Semiconductor Reticle Management Using Discrete Event Simulation". In *Proceedings of the 2007 Summer Computer Simulation Conference*, 593–600. San Diego, California: Society for Computer Simulation International.

Danilczyk, W., Y. L. Sun, and H. He. 2021. "Smart Grid Anomaly Detection using a Deep Learning Digital Twin". In *2020 52nd North American Power Symposium*, 1–6. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Eckhart, M., and A. Ekelhart. 2018. "Towards Security-Aware Virtual Environments for Digital Twins". In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security*, 61–72. New York, New York: ACM.

Eckhart, M., and A. Ekelhart. 2019. "Digital Twins for Cyber-physical Systems Security: State of the Art and Outlook". *Security and Quality in Cyber-Physical Systems Engineering*:383–412.

Eisenstein, Paul 2017. "European Car Plants Halted by WannaCry Ransomware Attack". https://www.nbcnews.com/business/autos/european-car-plants-halted-wannacry-ransomware-attack-n759496, accessed April 2023.

Garcia-Teodoro, P., J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez. 2009. "Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges". *Computers & Security* 28(1-2):18–28.

Hubballi, N., and V. Suryanarayanan. 2014. "False Alarm Minimization Techniques in Signature-based Intrusion Detection Systems: A Survey". *Computer Communications* 49:1–17.

IBM 2022. "Cost of a Data Breach 2022 Report". https://www.ibm.com/downloads/cas/E3G5JMBP, accessed May 2023.

Ioulianou, P., V. Vasilakis, I. Moscholios, and M. Logothetis. 2018. "A Signature-based Intrusion Detection System for the Internet of Things". *Information and Communication Technology Form*:11–13.

Jiang, Y., S. Yin, K. Li, H. Luo, and O. Kaynak. 2021. "Industrial Applications of Digital Twins". *Philosophical Transactions of the Royal Society A* 379(2207):20200360.

Kelley, Michael 2013. "The Stuxnet Attack On Iran's Nuclear Plant Was 'Far More Dangerous' Than Previously Thought". https://www.businessinsider.com/stuxnet-was-far-more-dangerous-than-previous-thought-2013-11, accessed May 2023.

Koroniotis, N., N. Moustafa, E. Sitnikova, and B. Turnbull. 2019. "Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-iot dataset". *Future Generation Computer Systems* 100:779–796.

Li, L., B. Lei, and C. Mao. 2022. "Digital Twin in Smart Manufacturing". *Journal of Industrial Information Integration* 26:100289.

Link Labs IoT Technology 2023. "Cost-Effective Solution for End-to-End Manufacturing Visibility". https://www.link-labs.com/hubfs/Link-Labs-AirFinder-RFID-for-Manufacturing.pdf, accessed May 2023.

Misra, S., and N. Saha. 2019. "Detour: Dynamic Task Offloading in Software-defined Fog for IoT Applications". *IEEE Journal on Selected Areas in Communications* 37(5):1159–1166.

Negri, E., L. Fumagalli, and M. Macchi. 2017. "A Review of the Roles of Digital Twin in CPS-based Production Systems". *Procedia Manufacturing* 11:939–948.

NOKIA Solutions 2023. "Accelerating Industry 4.0 digitalization and innovation". https://www.nokia.com/industry-4-0/, accessed May 2023.

United States Department of Homeland Security 2016. "ICS-CERT Year in Review 2016". https://www.cisa.gov/sites/default/files/Annual_Reports/Year_in_Review_FY2016_Final_S508C.pdf, accessed April 2023.

Park, S., J. Fowler, M. Carlyle, and M. Hickie. 1999. "Assessment of Potential Gains in Productivity Due to Proactive Reticle Management Using Discrete Event Simulation". In *Proceedings of the 1999 Winter Simulation Conference*, edited by P. Farrington, H. B. Nembhard, G. Evans, and D. Sturrock, 856–864. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Pedrielli, G., S. Chandrasekhar, D. Robert, R. Aditya, B. David, J. Mani, B. Eric, and G. Joseph. 2023. "DTFab: A Digital Twin based Approach for Optimal Reticle Management in Semiconductor Photolithography". *Journal of System Science and System Engineering, to appear*.

Sahal, R., S. H. Alsamhi, J. G. Breslin, K. N. Brown, and M. I. Ali. 2021. "Digital Twins Collaboration for Automatic Erratic Operational Data Detection in Industry 4.0". *Applied Sciences* 11(7):3186.

Schleich, B., N. Anwer, L. Mathieu, and S. Wartzack. 2017. "Shaping the Digital Twin for Design and Production Engineering". *CIRP Annals* 66(1):141–144.

Sturm, L. D., C. B. Williams, J. A. Camelio, J. White, and R. Parker. 2014. "Cyber-physical Vulnerabilities in Additive Manufacturing Systems". In *2014 International Solid Freeform Fabrication Symposium*. University of Texas at Austin.

Suhail, S., S. Zeadally, R. Jurdak, R. Hussain, R. Matulevičius, and D. Svetinovic. 2022. "Security Attacks and Solutions for Digital Twins". *arXiv preprint arXiv:2202.12501*.

Technology and Services Industry Association 2022. "The State of Field Services 2022". https://www.fieldtechnologiesonline.com/doc/the-state-of-field-services-0003, accessed May 2023.

Turner, H., J. White, J. A. Camelio, C. Williams, B. Amos, and R. Parker. 2015. "Bad Parts: Are Our Manufacturing Systems at Risk of Silent Cyberattacks?". *IEEE Security Privacy* 13(3):40–47.

Xu, Q., S. Ali, and T. Yue. 2021, April. "Digital Twin-based Anomaly Detection in Cyber-physical Systems". In *14th IEEE Conference on Software Testing*, 205–216. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Yampolskiy, M., T. R. Andel, J. T. McDonald, W. B. Glisson, and A. Yasinsac. 2015. "Towards Security of Additive Layer Manufacturing". *arXiv preprint arXiv:1602.07536*.

Zeltmann, S. E., N. Gupta, N. G. Tsoutsos, M. Maniatakos, J. Rajendran, and R. Karri. 2016. "Manufacturing and Security Challenges in 3D Printing". *The Journal of The Minerals, Metals Materials Society* 68(7):1872–1881.

Zhao, T., E. Foo, and H. Tian. 2022. "A Digital Twin Framework for Cyber Security in Cyber-Physical Systems". *arXiv preprint arXiv:2204.13859*.

AUTHOR BIOGRAPHIES

ROBERT DODGE is a Ph.D. student enrolled in the industrial engineering program at Arizona State University. He serves as a research assistant working under Dr. Giulia Pedrielli. His research focuses on stochastic simulation and optimization. His email is rwdodge@asu.edu

GIULIA PEDRIELLI received her Ph.D. in Mechanical Engineering from Politecnico di Milano, Italy in 2013. She pursued a Post-Doc at National University of Singapore from 2014-2016. She is currently Associate Professor for the School of Computing and Augmented Intelligence (SCAI) in Arizona State University. Her research is in the field of stochastic simulation and simulation optimization with a particular interest Bayesian Optimization. Her application areas span from logistics and supply chain, manufacturing, autonomous vehicles and bio-productions. Her email is gpedriel@asu.edu.

PETAR JEVTIĆ recieved his Ph.D. in Economics from Università degli Studi di Torino, Italy in 2013. He is currently an Assistant Professor for the School of Mathematical and Statistical Sciences at Arizona State University. His research involves modeling human mortality in a dynamic context, at the cohort level, the population level, and the multi-population level. He also works on various topic related to risk modeling, in particular cyber risk and climate risk, as well as classical topics of and property and casualty insurance, using predictive analytics tools and spatial modeling. His email is Petar.Jevtic@asu.edu