Optimal Auction Design with Malicious Sellers

Swastik Brahma
Department of Computer Science
Tennessee State University
Nashville, TN, USA
sbrahma@tnstate.edu

Laurent Njilla
Cyber Assurance Branch
Air Force Research Laboratory
Rome, NY, USA
laurent.njilla@us.af.mil

Satyaki Nan
Department of Computer Science
Tennessee State University
Nashville, TN, USA
snan@tnstate.edu

***Abstract*— The presence of hardware Trojans in Integrated Circuits (ICs) poses a serious security concern for the Internet-of-Things (IoT) and the mitigation of such threats has started to attract growing attention. This paper addresses the problem by designing a novel optimal auction mechanism that can optimize the utility of a buyer (system designer) acquiring ICs in a market where the sellers (manufacturers) can maliciously insert Trojans in sold ICs. The paper also proposes, as an integral component of our designed auction mechanism, the concept of redundant acquisition of ICs from multiple sellers to obtain reliability in terms of the buyer's system operation goals and enhance his utility as well as characterizes the optimal amount of redundancy the buyer should have in the acquisition process. The proposed auction mechanism can adapt to the imperfections in the process of testing acquired ICs. The optimal fine that should be imposed on a malicious seller in the auction upon detecting a Trojan in his sold IC is also characterized. Numerical results are presented to gain important insights into the proposed auction mechanism.**

***Index Terms*—IoT security, Hardware Trojans, Auction design.**

I. Introduction

Hardware Trojans are malicious modifications of the circuitry of Integrated Circuits (ICs). Such threats pose a serious security concern [1] for the Internet-of-Things (IoT) and can result in severe consequences such as leakage of sensitive information [1], [2], including leakage of sensed information [3], degradation of system performance [4], and even complete failure in achieving operational goals [5], [6], including those of medical IoT [7] systems.

The primary technique that past work [8]–[11] has focused on for mitigating threats from hardware Trojans is the development of testing strategies that can check for the presence of hardware Trojans in acquired ICs. For example, in [11], the authors have designed test patterns that can generate noticeable differences between the power profile of a genuine IC and that of an IC containing Trojans, but the effectiveness of their proposed technique is limited in terms of the manufacturing processes, behaviors and sizes of the Trojans. Since it can be prohibitive to exhaustively test an acquired IC against all possible Trojan types, the works in [12]–[16] develop techniques using game theory [17] that can intelligently determine which Trojan types should an acquired IC be tested for against a strategic malicious manufacturer. Specifically, [14] presents a two-person Trojan insertion-testing game, but investigates the equilibrium of the game for an example scenario. The game theoretic works in [12], [13] rely on software-based techniques for investigating testing strategies against a strategic malicious manufacturer at equilibrium. Analytical characterizations of equilibrium strategies in closed-forms for Trojan insertion-testing games in various scenarios can be found in [15], [16].

In contrast to past work, which has been limited to the development of testing strategies, with some works focusing on the development of game theoretic testing strategies [12]–[16], in this paper, we employ *Mechanism Design* [18], [19] (i.e., *reverse game theory*) to design a novel auction mechanism that can optimize the utility of a buyer acquiring IC(s) in a market environment where the sellers can act in a malicious manner by inserting Trojans in sold ICs. To the best of our knowledge, this is the first work that adopts a *mechanism design* perspective to address the problem of hardware Trojans. Further, to provide reliability to the buyer in terms of achieving his system operation goals and enhance his utility in such an environment, a problem which, to the best of our knowledge, has also not been addressed by past work, we propose the concept of using *redundancy* as an integral component of our designed auction mechanism. Specifically, the main contributions of our paper are as follows:

- We propose a novel optimal auction mechanism that can optimize the utility of a buyer who acquires IC(s) in a market environment where the sellers (manufacturers) can act in a malicious manner by inserting Trojans in sold ICs.
- In our designed auction mechanism, we introduce the novel concept of the buyer redundantly acquiring IC(s) from multiple sellers to obtain reliability in terms of successful system operation and enhance his utility. We also characterize the optimal amount of redundancy the buyer should have in the acquisition of IC(s).
- Our proposed auction mechanism considers the imperfections of the process of testing acquired ICs to check for the presence of Trojans and adapts accordingly.
- Numerical results are presented to gain important insights into our proposed auction mechanism and show its performance advantages.

The rest of the paper is organized as follows. Section II formulates the optimal auction design problem in the presence of malicious sellers. Section III analyzes the problem and presents the design of the optimal auction mechanism. Section IV provides numerical results to gain important insights. Finally, Section V concludes the paper.

This work was supported by the NSF under Grant CCF-2047701.
DISTRIBUTION A. Approved for public release. Distribution unlimited. Case Number AFRL-2021-3034. Dated 08 Sep 2021.

978-1-6654-1647-4/22/$31.00 ©2022 IEEE

Notation	Description
α	Probability of a seller being malicious
p	Probability of a malicious seller inserting a Trojan in his sold IC
P_d	Probability of detecting an inserted Trojan
V_i	Benefit acquired by malicious seller i from a successfully inserted Trojan
G_i	Fine imposed on malicious seller i upon detecting a Trojan in his sold IC
B^S	Benefit acquired by the buyer from using a Trojan-free IC
t_i	Bid of seller i, where $t_i \in [a_i, b_i]$
$f_i(.)$	Probability density function (PDF) of t_i
$F_i(.)$	Cumulative distribution function (CDF) of t_i
m_i	Payment to be made to seller i
q_i	Selection state of seller i, where $q_i \in \{0,1\}$
d	Number of redundantly selected sellers

TABLE I
NOTATIONS USED

II. FORMULATION OF THE AUCTION DESIGN PROBLEM

In this section, we formulate the problem of optimal auction design in the presence of malicious sellers. Specifically, consider a buyer (system designer) who conducts an auction to acquire IC(s) from a set $\{1, \cdots, N\}$ comprised of N sellers (manufacturers). Suppose that seller $i \in \{1, \cdots, N\}$ has the (true) valuation $t_i \in [a_i, b_i]$ for his IC, where t_i is considered as a random variable with the probability density function (PDF) $f_i(\cdot)$, and a_i and b_i are the lowest and highest possible valuations of seller i for his IC, respectively. Suppose also that every seller can be malicious in nature with a probability α and that every malicious seller inserts a Trojan into his sold IC with a probability p. Consider that the sellers announce their valuations of their ICs (as their bids) to the *buyer*, who, upon getting a vector of bids $\mathbf{t} = (t_1, \cdots, t_N)$ from the sellers, suppose uses the functions:

- $\mathbf{q}(\mathbf{t}) = \big(q_1(\mathbf{t}), \cdots, q_N(\mathbf{t})\big)$ to determine which seller(s) to acquire IC(s) from, where $q_i(\mathbf{t}) \in \{0, 1\}$ assumes Boolean values to denote whether seller i is selected, and
- $\mathbf{m}(\mathbf{t}) = \big(m_1(\mathbf{t}), \cdots, m_N(\mathbf{t})\big)$ to determine the payments to be made to the sellers, where $m_i(\mathbf{t})$ is the payment to be made to seller i.

We consider the fact that the sellers, to gain undue advantages, can announce *falsified* valuations as their bids. Further, we consider that the buyer, after acquiring an IC from a seller, *tests* the IC to check for the presence of Trojans. To model the imperfections of the testing process, we consider that a Trojan inserted in an IC is detected by the buyer with a probability P_d. We also consider that, upon detecting a Trojan in an IC bought from seller i, the seller is imposed a fine G_i (and the buyer, of course, does not put to use such an IC). If, however, an inserted Trojan remains undetected in an IC bought from seller i, and the buyer uses (installs) such an IC, we consider that the inserted Trojan accomplishes its goal providing the seller a benefit V_i (while negatively impacting the buyer's utility by an amount V_i reflecting the damage caused to the buyer).

Further, in our auction mechanism, we propose the novel concept of the buyer *redundantly* buying the required IC from multiple sellers (specifically, from $d \leq N$ sellers) to enhance the chances of the buyer acquiring a Trojan-free IC (with the buyer getting a benefit B^S upon being able to use a Trojan-free IC). Table I summarizes our notations used.

The goal of our auction design problem is to design the functions $\mathbf{q}(\mathbf{t})$ and $\mathbf{m}(\mathbf{t})$ such that the expected utility of the buyer from using the auction mechanism is maximized while satisfying certain constraints. Next, we describe the expected utilities of the buyer and the sellers from the aforementioned auction mechanism and then describe the constraints that need to be satisfied along with the formulation of the auction design problem as an optimization problem.

A. Expected Utilities

The expected utility ($U^B(\mathbf{m}, \mathbf{q})$) of the buyer from the auction mechanism is

$$U^B(\mathbf{m}, \mathbf{q}) = B^S\big[1 - (\alpha p)^d\big] \left[\frac{1 - \alpha p}{1 - \alpha p P_d}\right] + \int_T \Bigg[\alpha p P_d \sum_{i=1}^{N} q_i(\mathbf{t}) G_i - \alpha p (1 - P_d) k \sum_{i=1}^{N} V_i q_i(\mathbf{t}) - \sum_{i=1}^{N} m_i(\mathbf{t})\Bigg] f(\mathbf{t}) d\mathbf{t} \quad (1)$$

where, $T = [a_1, b_1] \times \cdots \times [a_N, b_N]$ denotes the set of all possible combinations of the sellers' valuations, $f(\mathbf{t}) = \prod_{i=1}^{N} f_i(t_i)$ is the joint density function on T for the vector of valuations $\mathbf{t} = (t_1, \cdots, t_N)$, and $dt = dt_1 \cdots dt_N$. Further, considering that the buyer selects an IC to be used from the acquired ones in which a Trojan was not detected via conducted tests (with the tests themselves being error-prone in nature) with uniform probability, in (1), we have

$$k = \frac{1}{d(1 - \alpha p P_d)} \quad (2)$$

which is the expected probability with which a seller's IC is installed by the buyer given that the buyer buys an IC from the seller and that the bought IC tests negative for the presence of a Trojan. Further, accordingly, $\big[1 - (\alpha p)^d\big] \left[\frac{1 - \alpha p}{1 - \alpha p P_d}\right]$ in (1) is the probability of the buyer using a Trojan-free IC from among the d ICs acquired from d different sellers.

The expected utility ($U_i^S(m_i, q_i, t_i)$) of seller i from the auction mechanism, having a true valuation $t_i \in [a_i, b_i]$ regarding his IC, is

$$U_i^S(m_i, q_i, t_i) = \int_{T_{-i}} \Big[m_i(\mathbf{t}) - q_i(\mathbf{t})\Big\{t_i - \alpha p\big((1 - P_d) k V_i - P_d G_i\big)\Big\}\Big] f_{-i}(\mathbf{t}_{-i}) d\mathbf{t}_{-i} \quad (3)$$

where, $T_{-i} = [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_N, b_N]$ denotes the set of all possible combinations of the sellers' valuations other than i, $f_{-i}(\mathbf{t}_{-i}) = \prod_{j \in \{1, \cdots, N\}, j \neq i} f_j(t_j)$ is the joint density function on T_{-i} for the vector of valuations $\mathbf{t}_{-i} = (t_1, \cdots, t_{i-1}, t_{i+1}, \cdots, t_N)$, $dt_{-i} = dt_1 \cdots dt_{i-1} dt_{i+1} \cdots dt_N$, and k follows its definition in (2).

Further, we consider that every seller i, having a true valuation t_i, may send a falsified bid w_i hoping to make an undue profit from the auction mechanism, which would make the expected utility of the seller to be

$$\tilde{U}_i^S(m_i, q_i, w_i) = \int_{T_{-i}} \Big[m_i(w_i, \mathbf{t}_{-i}) - q_i(w_i, \mathbf{t}_{-i})\Big\{t_i - \alpha p\big((1 - P_d) k V_i - P_d G_i\big)\Big\}\Big] f_{-i}(\mathbf{t}_{-i}) d\mathbf{t}_{-i} \quad (4)$$

B. Optimization Problem

The auction design problem can be formulated as the following optimization problem:

$$\max_{\mathbf{m},\mathbf{q}} \; U^B(\mathbf{m},\mathbf{q})$$

Subject to:

$$U_i^S(m_i,q_i,t_i) \geq 0, \quad \forall i \in \{1,\cdots,N\}, \forall t_i \in [a_i,b_i] \tag{5a}$$

$$U_i^S(m_i,q_i,t_i) \geq \tilde{U}_i^S(m_i,q_i,w_i), \quad \forall i \in \{1,\cdots,N\}, \forall t_i, w_i \in [a_i,b_i] \tag{5b}$$

$$q_i(\mathbf{t}) \in \{0,1\}, \quad \forall i \in \{1,\cdots,N\} \tag{5c}$$

$$\sum_{i=1}^{N} q_i(\mathbf{t}) = d \tag{5d}$$

Below we describe each constraint in detail:

- *Individual-Rationality constraint* (5a): We consider that the buyer cannot force a seller to participate in an auction. If a seller did not participate in the auction, clearly his utility would be zero. Thus, to rationalize participation of sellers in the auction, the utility of every seller must be greater than or equal to zero.
- *Incentive-Compatibility constraint* (5b): We consider that the buyer can not prevent any seller from providing a falsified valuation as his bid if the seller expects to gain from lying. Thus, to prevent sellers from lying about their valuations of their ICs, honest reporting of valuations during the bidding process must form a Nash equilibrium in the auction game.
- *Selection Parameter constraint* (5c): The selection parameter of every seller is a Boolean variable.
- *Redundancy constraint* (5d): The buyer selects d sellers for buying to have the desired amount of redundancy.

Next, we analyze the above auction design problem.

III. Analysis of the Auction Design Problem

Let us denote the expected value of the selection parameter of seller i, viz. $q_i(t_i,\mathbf{t}_{-i})$, for a given t_i, as

$$Q_i(q_i,t_i) = \int_{T_{-i}} q_i(t_i,\mathbf{t}_{-i}) f_{-i}(\mathbf{t}_{-i}) d\mathbf{t}_{-i} \tag{6}$$

Our first result is a simplified characterization of the incentive-compatibility constraint (5b) presented in Section II-B.

Lemma 1. *The incentive-compatibility constraint* (5b) *holds only if the following two conditions hold* $\forall i \in \{1,\cdots,N\}$*:*

1. *If* $t_i \leq w_i$, *then* $Q_i(q_i,w_i) \leq Q_i(q_i,t_i)$ (7a)
2. $U_i^S(m_i,q_i,t_i) = U_i^S(m_i,q_i,b_i) + \int_{t_i}^{b_i} Q_i(q_i,w_i)dw_i$ (7b)

Proof. Consider $t_i, w_i \in [a_i,b_i]$ with $t_i \leq w_i$. Also, suppose that, while t_i is the true valuation of seller i for his IC, he sends the falsified valuation w_i as his bid. In this case, seller i's utility (4) can be rewritten as

$$\int_{T_{-i}} \Big[m_i(w_i,\mathbf{t}_{-i}) - q_i(w_i,\mathbf{t}_{-i}) \Big[w_i - \alpha p \big\{ (1-P_d) kV_i - P_d G_i \big\} \Big] + q_i(w_i,\mathbf{t}_{-i})(w_i - t_i) \Big] f_{-i}(\mathbf{t}_{-i}) d\mathbf{t}_{-i} \tag{8}$$

$$= U_i^S(m_i,q_i,w_i) + (w_i - t_i) Q_i(q_i,w_i) \quad \text{(using (6))} \tag{9}$$

The incentive-compatibility constraint (5b) states that the expected utility of every seller i from reporting his true valuation must be greater than or equal to his expected utility from reporting a falsified valuation as his bid. Thus, we must have

$$U_i^S(m_i,q_i,t_i) \geq U_i^S(m_i,q_i,w_i) + (w_i - t_i) Q_i(q_i,w_i) \tag{10}$$

which implies that

$$(w_i - t_i) Q_i(q_i,w_i) \leq U_i^S(m_i,q_i,t_i) - U_i^S(m_i,q_i,w_i) \tag{11}$$

Similarly, considering w_i to be the true valuation of seller i and t_i to be a falsified valuation, the incentive-compatibility constraint implies

$$(w_i - t_i) Q_i(q_i,t_i) \geq U_i^S(m_i,q_i,t_i) - U_i^S(m_i,q_i,w_i) \tag{12}$$

From (11) and (12), we get

$$(w_i - t_i) Q_i(q_i,w_i) \leq U_i^S(m_i,q_i,t_i) - U_i^S(m_i,q_i,w_i) \leq (w_i - t_i) Q_i(q_i,t_i) \tag{13}$$

Clearly, (13) implies (7a). Further, defining $\delta = w_i - t_i$, we can rewrite the inequalities in (13) for $\delta \to 0$ as

$$Q_i(q_i,w_i)\delta \leq U_i^S(m_i,q_i,w_i-\delta) - U_i^S(m_i,q_i,w_i) \leq Q_i(q_i,w_i-\delta)\delta \tag{14}$$

Thus, $Q_i(q_i,w_i)$ is a decreasing function of w_i, and thus Riemann integrable, based on which we get

$$U_i^S(m_i,q_i,t_i) = U_i^S(m_i,q_i,b_i) + \int_{t_i}^{b_i} Q_i(q_i,w_i)dw_i \tag{15}$$

which proves (7b). This proves the lemma. □

Based on Lemma 1, the optimization problem in Section II-B can be simplified as given in the following theorem.

Theorem 1. *For* $(\mathbf{m},\mathbf{q})$ *to represent an optimal auction mechanism,* $\mathbf{q}(\mathbf{t})$ *should maximize*

$$B^S\big[1-(\alpha p)^d\big]\Big[\frac{1-\alpha p}{1-\alpha p P_d}\Big] - \int_T \sum_{i=1}^{N} \Big[t_i + \frac{F_i(t_i)}{f_i(t_i)} \Big] q_i(\mathbf{t}) f(\mathbf{t}) d(\mathbf{t}) \tag{16}$$

subject to constraints (5c) and (5d), and the payment to seller i *should be given by*

$$m_i(\mathbf{t}) = q_i(\mathbf{t})\big[t_i + \alpha p\{P_d G_i - (1-P_d)kV_i\}\big] + \int_{t_i}^{b_i} q_i(w_i,\mathbf{t}_{-i}) dw_i \tag{17}$$

Proof. The buyer's expected utility (1) can be rewritten as

$$U^B(\mathbf{m},\mathbf{q}) = B^S\left[1-(\alpha p)^d\right]\left[\frac{1-\alpha p}{1-\alpha p P_d}\right] + \sum_{i=1}^{N} \int_T \Big[-m_i(\mathbf{t}) + q_i(\mathbf{t}) \Big\{ t_i - \alpha p \big((1-P_d)kV_i - P_d G_i \big) \Big\} \Big] f(\mathbf{t}) d\mathbf{t} - \sum_{i=1}^{N} \int_T t_i q_i(\mathbf{t}) f(\mathbf{t}) d\mathbf{t} \tag{18}$$

Now, we have

$$\int_T \Big[-m_i(\mathbf{t}) + q_i(\mathbf{t}) \Big\{ t_i - \alpha p \big((1-P_d) k V_i - P_d G_i \big) \Big\} \Big] f(\mathbf{t}) d\mathbf{t} \tag{19}$$

$$= -\int_{a_i}^{b_i} U_i^S(m_i, q_i, t_i) f_i(t_i) dt_i \tag{20}$$

$$= -\int_{a_i}^{b_i} \Big[U_i^S(m_i, q_i, b_i) + \int_{t_i}^{b_i} Q_i(q_i, w_i) dw_i \Big] f_i(t_i) dt_i \tag{21}$$

$$= -U_i^S(m_i, q_i, b_i) - \int_{a_i}^{b_i} \int_{a_i}^{w_i} Q_i(q_i, w_i) dw_i f_i(t_i) dt_i \tag{22}$$

$$= -U_i^S(m_i, q_i, b_i) - \int_{a_i}^{b_i} F_i(w_i) Q_i(q_i, w_i) dw_i \tag{23}$$

$$= -U_i^S(m_i, q_i, b_i) - \int_T \frac{F_i(t_i)}{f_i(t_i)} q_i(\mathbf{t}) f(\mathbf{t}) d\mathbf{t} \tag{24}$$

Substituting (24) into (18), we get,

$$U^B(\mathbf{m}, \mathbf{q}) = B^S \left[1 - (\alpha p)^d \right] \left[\frac{1-\alpha p}{1-\alpha p P_d} \right] - \int_T \sum_{i=1}^{N} \Big[t_i + \frac{F_i(t_i)}{f_i(t_i)} \Big] q_i(\mathbf{t}) f(\mathbf{t}) d(\mathbf{t}) - \sum_{i=1}^{N} U_i^S(m_i, q_i, b_i) \tag{25}$$

In (25), $\mathbf{m}(\mathbf{t})$ only appears in the last term of the buyer's utility function. Also, from the individual-rationality constraint (5a), we know that for every seller i, we must have $U_i^S(m_i, q_i, b_i) \geq 0$. Thus, the best possible value of the last term of (25) can be obtained, which is zero (since the buyer seeks to maximize his utility), as well as the individual-rationality constraint can be satisfied, by having $U_i^S(m_i, q_i, b_i) = 0$, $\forall i \in \{1, \cdots, N\}$, which implies, using (7b),

$$U_i^S(m_i, q_i, t_i) - \int_{t_i}^{b_i} Q_i(q_i, w_i) dw_i = 0 \tag{26}$$

Substituting (3) and (6) into (26), we get (17) with the utility of the buyer thereby becoming (16). This proves the theorem. □

A. Determination of Auction Outcome

We now describe how to determine the outcome of the optimal auction mechanism characterized in Theorem 1. In particular, we describe how to find the optimal set of sellers, their payments, and the optimal amount of fine that should be imposed on a seller upon finding a Trojan in his sold IC. We also discuss the optimization of d to characterize the optimal amount of redundancy the buyer should opt for.

1) Optimal seller selection for a given d: Given valuation $t_i \in [a_i, b_i]$ of seller i, let us define

$$\eta_i(t_i) = t_i + \frac{F_i(t_i)}{f_i(t_i)} \tag{27}$$

We refer to $\eta_i(t_i)$ as the *virtual valuation* of seller i. Now, it can be noted that the buyer's expected utility (16) in the optimal auction mechanism is maximized if $\mathbf{q}(\mathbf{t})$ is such that it maximizes

$$B^S [1 - (\alpha p)^d] \left[\frac{1-\alpha p}{1-\alpha p P_d} \right] - \sum_{i=1}^{N} \eta_i(t_i) q_i(\mathbf{t}) \tag{28}$$

subject to constraints (5c) and (5d) for all $\mathbf{t} \in T$. To maximize (28) for a given d (i.e., for a given number of sellers to be selected), clearly, upon receiving a vector of bids $\mathbf{t} = (t_1, \cdots, t_N)$ from the sellers, the buyer should first calculate the virtual valuations of the sellers (using (27)), number the sellers in non-decreasing order of their virtual valuations (such that $\eta_1(t_1) \leq \eta_2(t_2) \leq \cdots \leq \eta_N(t_N)$), and then select the d sellers having the lowest virtual valuations, i.e., choose $q_i(\mathbf{t}) = 1$ (if $i \in [1, d]$) and $q_i(\mathbf{t}) = 0$ (otherwise).

2) Optimal seller section while optimizing d: (28) can clearly be optimized by optimally selecting d. To find optimal d, we first prove some properties of the first term of (28) in the following lemma.

LEMMA 2. *The first term of* (28)*, viz.* $B^S[1-(\alpha p)^d][\frac{1-\alpha p}{1-\alpha p P_d}]$*, is an increasing function of* d*. The rate of increase of the first term of* (28) *is non-increasing with* d*.*

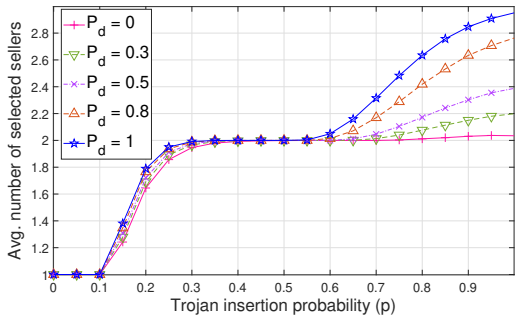
Proof. Let us denote the first term of (28) as

$$\xi(d) = B^S \left[1 - (\alpha p)^d \right] \left[\frac{1-\alpha p}{1-\alpha p P_d} \right] \tag{29}$$

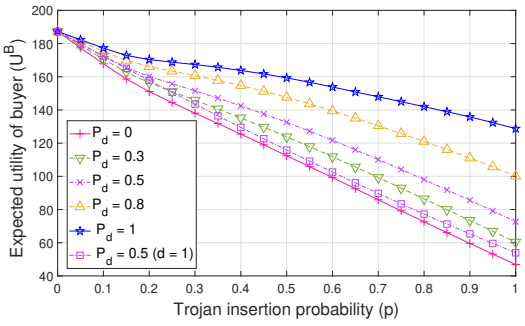
Now, since $\alpha p \in [0, 1]$, we have $\xi(d = \kappa + 1) - \xi(d = \kappa) = B^S[\frac{1-\alpha p}{1-\alpha p P_d}][(\alpha p)^\kappa - (\alpha p)^{\kappa+1}] \geq 0$, showing that $\xi(d)$ is an increasing function of d. Again, since $\alpha p \in [0, 1]$, we have $\xi(d = \kappa+1) - \xi(d = \kappa) \geq \big\{ \xi(d = \kappa+1) - \xi(d = \kappa) \big\} (\alpha p) = \xi(d = \kappa + 2) - \xi(d = \kappa + 1)$, which shows that the rate of increase of $\xi(d)$ is non-increasing with d. □

Based on Lemma 2, upon receiving a vector of bids $\mathbf{t} = (t_1, \cdots, t_N)$, to find the optimal value of d, the buyer should compute the virtual valuation $\eta_i(t_i)$ (27) for every seller i, number the sellers in non-decreasing order of their virtual valuations $\big($such that $\eta_1(t_1) \leq \eta_2(t_2) \leq \cdots \leq \eta_N(t_N)\big)$, and then, to maximize (28), choose the *largest* value of $\kappa \in [1, N]$ that satisfies the condition $\eta_\kappa(t_\kappa) \leq \xi(d = \kappa) - \xi(d = \kappa - 1)$ as d^* (i.e., as the optimal value of d), where $\xi(\cdot)$ follows (29). Clearly, the optimal set of d^* sellers (that would maximize (28)) corresponds to the buyer choosing $q_i(\mathbf{t}) = 1$ (if $i \in [1, d^*]$) and $q_i(\mathbf{t}) = 0$ (otherwise). If, however, $\eta_1(t_1) > \xi(d = 1) - \xi(d = 0)$, then there clearly does not exist any value of $\kappa \in [1, N]$ that satisfies the aforementioned condition and, in such a scenario, the buyer should choose $d^* = 0$ (i.e., the buyer should not buy the IC from any seller). Finally, it should be noted that the above procedure for determining d^* and the optimal set of d^* sellers requires sorting the sellers in non-decreasing order of their virtual valuations, thereby having a time complexity of $\mathcal{O}(N \log N)$.

3) Determination of Sellers' Payments: The payments to be made to the sellers can be computed using (17). First, it can be noted from (17) that sellers who are not selected by the buyer do not receive any payment, since if seller i is not selected with a bid $t_i \in [a_i, b_i]$, then $q_i(w_i, \mathbf{t}_{-i}) = 0$, $\forall w_i \in [t_i, b_i]$. The difficulty of computing the payment of seller i who is selected by the buyer lies in calculating the integral term of (17). To compute the integral term, it can be noted that seller i who is selected with a bid $t_i \in [a_i, b_i]$ could have still been selected with a higher bid until his bid exceeds a certain cutoff bid,

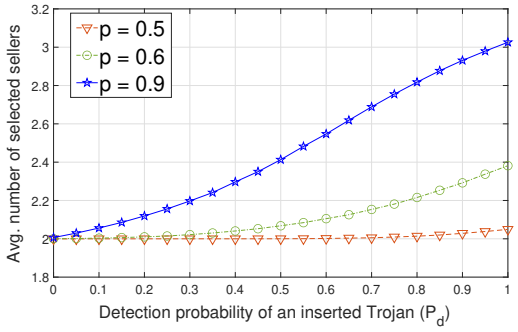


(a)

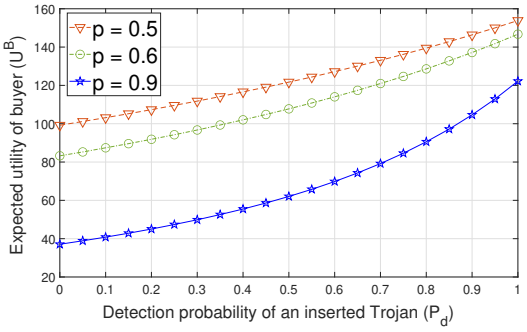


(b)

Fig. 1. Impact of Trojan insertion probability (p) on auction outcome.



(a)



(b)

Fig. 2. Impact of the detection probability (P_d) on auction outcome.

say $\sigma_i \in [t_i, b_i]$. In other words, we have $q_i(w_i, \mathbf{t}_{-i}) = 1$, $\forall w_i \in [t_i, \sigma_i]$, and $q_i(w_i, \mathbf{t}_{-i}) = 0$, $\forall w_i \in (\sigma_i, b_i]$. Thus, the payment (17) of seller i who is selected by the buyer becomes $m_i(\mathbf{t}) = t_i + \alpha p\{P_d G_i - (1-P_d)kV_i\} + (\sigma_i - t_i) = \sigma_i + \alpha p\{P_d G_i - (1-P_d)kV_i\}$. The cutoff bid σ_i for seller i can be determined in a computationally efficient manner by using the bisection method [20] to iteratively narrow down the highest possible winning bid of seller i in the range $[t_i, b_i]$.

Further, noting that (17) is a function of the fine (G_i) imposed on seller i upon detecting a Trojan in his sold IC leads us to the following important remark regarding the optimal fine that should be charged.

REMARK 1. *The optimal fine that should be imposed on seller i upon detecting a Trojan in his sold IC can be found by setting the term $\{P_d G_i - (1-P_d)kV_i\}$ in (17) to 0, which yields $G_i = \frac{1-P_d}{P_d} kV_i$. This results in the least amount of payment to be made to a selected seller while ensuring that the seller is unable to make a profit by sending a falsified bid.*

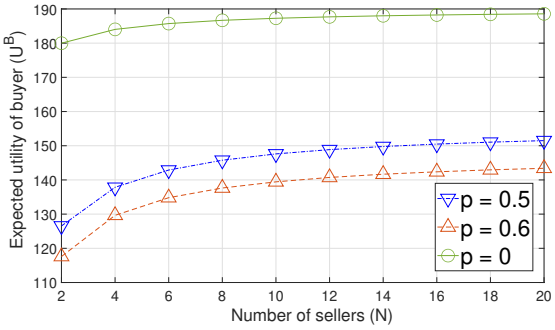
IV. NUMERICAL RESULTS

In this section, we provide numerical results to gain insights into the proposed auction mechanism and show its performance advantages. To obtain the results, the optimal number of redundant sellers along with the optimal set of sellers to be selected were determined using the procedure described in Section III-A2 for optimizing the buyer's expected utility in (16), and the payments of the selected sellers were determined using the procedure described in Section III-A3 for computing (17) with the fines set optimally based on Remark 1.

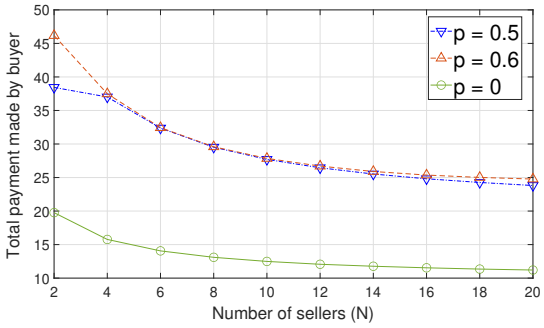
In Fig. 1, we show how the probability (p) with which a malicious seller inserts a Trojan into his sold IC impacts the outcome of the proposed auction mechanism. For the figure, we consider, $N = 10$, $B^S = 200$, $\alpha = 0.5$, and the valuations of all sellers for their ICs to be uniformly distributed over the range $[10, 25]$. As can be seen from Fig. 1(a), as p increases, for a given P_d, the average number of sellers redundantly selected by the buyer shows a non-decreasing trend. This happens to mitigate the impact of larger fractions of acquired ICs having Trojans in them with increasing p. As can also be noted from the figure, for a given p, the average number of sellers redundantly selected shows a non-decreasing trend with P_d. This is because *a higher P_d enhances the ability of the buyer to find a Trojan-free IC from among the acquired ones thereby rationalizing the buyer's investment in buying ICs from multiple sellers with a higher degree of redundancy with increasing P_d*. As can be noted from Fig. 1(b), and as is also intuitive, the expected utility of the buyer in the proposed auction mechanism decreases with increasing p for a given P_d. Further, as expected, for a given p, it can be noted from the figure that the expected utility of the buyer increases with P_d. Moreover, Fig. 1(b) also shows the expected utility of the buyer (for $P_d = 0.5$) with redundant acquisition of ICs from multiple sellers allowed as well as when its not (i.e., for $d = 1$). As can be seen, the expected utility of the buyer when redundant acquisition of ICs from multiple sellers is allowed is greater than or equal to that of the case where $d = 1$, which *shows the performance advantage* of redundantly acquiring ICs from multiple sellers in the proposed mechanism.

In Fig. 2, we show the impact of the detection probability (P_d) of an inserted Trojan on the outcome of the proposed auction mechanism. For the figure, we consider $N = 10$, $B^S = 200$, $\alpha = 0.6$, and the valuations of all sellers for

their ICs to be uniformly distributed over the range $[10, 25]$. As can be seen from Fig. 2(a), the average number of sellers redundantly selected by the buyer shows a non-decreasing trend with P_d (for a given Trojan insertion probability p) as well as shows a non-decreasing trend with p (for a given P_d) due to reasons discussed earlier for Fig. 1(a). Moreover, as expected, the expected utility of the buyer increases with P_d (for a given p) and decreases with p (for a given P_d) as noted earlier for Fig. 1(b). These observations corroborate those made for Fig. 1.



(a)



(b)

Fig. 3. Impact of the number of sellers (N) on auction outcome.

In Fig. 3, we show how the number of sellers (N) impacts the outcome of the proposed auction mechanism. For the figure, we consider $B^S = 200$, $\alpha = 0.5$, $P_d = 0.8$, and the valuations of all sellers for their ICs to be uniformly distributed over the range $[10, 25]$. As can be seen from Fig. 3(a), the expected utility of the buyer in the proposed auction mechanism shows an increasing trend with N, for a given probability of Trojan insertion (p), since with N the chances of finding sellers who can sell for lower payments as well as the number of redundant sellers that can be selected both increase. As expected, the expected utility of the buyer, for a given N, decreases with p. As can be seen from Fig. 3(b), and as is consistent with the observation made for Fig. 3(a), the total payment made by the buyer shows a decreasing trend with N. Moreover, as can be seen, for a given N, the total payment made by the buyer shows a non-decreasing trend with p since, as noted in Fig. 1(a), the number of redundantly selected sellers shows a non-decreasing trend with p, thereby making the total payment needed follow a non-decreasing trend.

V. Conclusion

This paper adopted a mechanism design perspective to design an optimal auction mechanism that can optimize the utility of a buyer acquiring ICs in a market where sellers can act in a malicious manner by inserting hardware Trojans in sold ICs. The presence of Trojans in constituent ICs poses a severe security threat to IoT. The proposed mechanism can optimally employ redundancy in the selection of sellers to obtain reliability and enhance the buyer's utility. Our auction mechanism can adapt to the imperfections of the process of testing acquired ICs. The optimal fine that should be imposed on a malicious seller upon detecting a Trojan in his sold IC was also characterized. Numerical results were presented to gain important insights into the proposed auction mechanism.

References

[1] O. Arias, J. Wurm, K. Hoang, and Y. Jin, "Privacy and security in internet of things and wearable devices," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 1, no. 2, pp. 99–109, 2015.

[2] S. Guo, J. Wang, Z. Chen, Y. Li, and Z. Lu, "Securing iot space via hardware trojan detection," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11 115–11 122, 2020.

[3] C. Dong, G. He, X. Liu, Y. Yang, and W. Guo, "A multi-layer hardware trojan protection framework for iot chips," *IEEE Access*, 2019.

[4] J. Dofe, J. Frey, and Q. Yu, "Hardware security assurance in emerging iot applications," in *IEEE Intl. Symp. on Circuits and Sys. (ISCAS)*, 2016.

[5] A. Nourian and S. Madnick, "A systems theoretic approach to the security threats in cyber physical systems applied to stuxnet," *IEEE Trans. on Dependable and Secure Comp.*, vol. 15, no. 1, pp. 2–13, 2018.

[6] S. Adee, "The hunt for the kill switch," *IEEE Spectrum*, vol. 45, no. 5, pp. 34–39, 2008.

[7] Z. Kahleifeh, S. Dinesh Kumar, and H. Thapliyal, "Hardware trojan detection in implantable medical devices using adiabatic computing," in *2018 IEEE Intl. Conf. on Rebooting Computing (ICRC)*, 2018, pp. 1–6.

[8] S. Bhunia, M. S. Hsiao, M. Banga, and S. Narasimhan, "Hardware trojan attacks: Threat analysis and countermeasures," *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1229–1247, 2014.

[9] M. Banga and M. S. Hsiao, "A region based approach for the identification of hardware trojans," in *IEEE International Workshop on Hardware-Oriented Security and Trust*, 2008, pp. 40–47.

[10] R. S. Chakraborty, S. Narasimhan, and S. Bhunia, "Hardware trojan: Threats and emerging solutions," in *IEEE International High Level Design Validation and Test Workshop*, 2009, pp. 166–171.

[11] D. Agrawal, S. Baktir, D. Karakoyunlu, P. Rohatgi, and B. Sunar, "Trojan detection using ic fingerprinting," in *IEEE Symposium on Security and Privacy*, 2007, pp. 296–310.

[12] C. A. Kamhoua, H. Zhao, M. Rodriguez, and K. A. Kwiat, "A game-theoretic approach for testing for hardware trojans," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 2, no. 3, pp. 199–210, 2016.

[13] J. Graf, W. Batchelor, S. Harper, R. Marlow, E. Carlisle, and P. Athanas, "A practical application of game theory to optimize selection of hardware trojan detection strategies," *Journal of Hardware and Systems Security*, vol. 4, no. 2, pp. 98–119, 2020.

[14] J. Graf, "Trust games: How game theory can guide the development of hardware trojan detection methods," in *IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2016, pp. 91–96.

[15] S. Brahma, S. Nan, and L. Njilla, "Strategic hardware trojan testing with hierarchical trojan types," in *55th Annual Conference on Information Sciences and Systems (CISS)*, 2021, pp. 1–6.

[16] S. Brahma, L. Njilla, and S. Nan, "Game theoretic hardware trojan testing under cost considerations," in *International Conference on Decision and Game Theory for Security*. Springer, 2021, pp. 251–270.

[17] D. Fudenberg and J. Tirole, *Game Theory, MIT Press*, 1991.

[18] N. Nisan and A. Ronen, "Algorithmic mechanism design," *Games and Economic behavior*, vol. 35, no. 1-2, pp. 166–196, 2001.

[19] R. B. Myerson, "Optimal auction design," *Mathematics of operations research*, vol. 6, no. 1, pp. 58–73, 1981.

[20] J. F. Epperson, *An introduction to numerical methods and analysis*, 2021, John Wiley & Sons.