Interdependent Security Games in the Stackelberg Style: How First-Mover Advantage Impacts Free-Riding and Security (Under-)Investment

Ziyuan Huang, Parinaz Naghizadeh, Mingyan Liu *

Abstract

Network games are commonly used to capture the strategic interactions among interconnected agents in simultaneous moves. The agents' actions in a Nash equilibrium must take into account the mutual dependencies connecting them, which is typically obtained by solving a set of fixed point equations. Stackelberg games, on the other hand, model the sequential moves between agents that are categorized as leaders and followers. The corresponding solution concept, the subgame perfect equilibrium, is typically obtained using backward induction. Both game forms enjoy very wide use in the (cyber)security literature, the network game often as a template to study security investment and externality – also referred to as the Interdependent Security (IDS) games – and the Stackelberg game as a formalism to model a variety of attacker-defender scenarios.

In this study we examine a model that combines both types of strategic reasoning: the interdependency as well as sequential moves. Specifically, we consider a scenario with a network of interconnected first movers (firms or defenders, whose security efforts and practices collectively determine the security posture of the eco-system) and one or more second movers, the attacker(s), who determine how much effort to exert on attacking the many potential targets. This gives rise to an equilibrium concept that embodies both types of equilibria mentioned above. We will examine how its existence and uniqueness conditions differ from that for a standard network game. Of particular interest are comparisons between the two game forms in terms of effort exerted by the defender(s) and the attacker(s), respectively, and the free-riding behavior among the defenders.

1 Introduction

Network games are commonly used to capture the strategic interactions among interconnected agents. As a special type of strategic game, network games highlight the critical role physical or logical connectivity can play in determining the outcome of social, technological, and economic interactions, see e.g., [7, 14, 15, 20, 22, 25, 26]; this is seen in many real-world scenarios such as trade, exchange and spread of information, and the marketing and adoption of technology, to name a few. This

*Ziyuan Huang and Mingyan Liu are with the Electrical and Computer Engineering Department, University of Michigan, Ann Arbor, MI 48109, USA; e-mail: {ziyuanh, mingyan} @umich.edu; Parinaz Naghizadeh is with the Integrated Systems Engineering & Electrical and Computer Engineering Departments, The Ohio State University, Columbus, OH 43210, USA; e-mail: naghizadeh.1@osu.edu.

framework is widely used in the context of local provision of public goods, see e.g., [6, 9, 22]. One of the primary goals of this line of research has been to identify how properties of the connectivity, i.e., the underlying interaction graph that represents mutual dependencies and influences among the agents, impact the equilibrium outcomes of such games in terms of their analysis, computation, as well as the conditions under which an equilibrium solution exists, is unique and/or is stable. Network games are almost always studied as a strategic, simultaneous-move game, with the notion of *Nash equilibrium (NE)* used to study their outcomes.

In the security context, public good provision games are often utilized as a framework to model strategic defense decisions; examples include airline security [19] and cybersecurity [11, 13, 21, 22, 29], where network games have often been referred to as interdependent security (IDS) games. These models are used to study the provision of, or investment in, security and defense mechanisms by a set of interconnected and interdependent entities (representing companies, agencies, organizations). Typically the impact of security investment or effort in such a model is captured through the agents' utility function, which is an increasing function of effort (or more precisely, some type of total weighted effort to reflect its dependency on other agents' effort); an adverse event or the existence of an attacker or adversary is not explicitly modeled, with the exception of the seminal work by Bracken and McGill [4, 5] (and studies that followed) on the two-stage modeling of the defense-assault application. Bracken and McGill also pioneered the rich field of bilevel programming, which has found broad applicability in adversarial security games.

In this paper, we are interested in understanding more explicitly the interaction between interconnected defenders and similarly interconnected attackers. The differentiation between the defenders and the attackers not only lies in the dependence relationship (e.g., a defender may benefit from a fellow defender's effort, but will suffer harm from an attacker's effort), but also in the sequentiality of their actions; that is, a (targeted) attack often takes into account some knowledge of the state of the defender. In practical terms, this assumption is supported by the observation that cyber criminals, in order to accomplish their attack goals, often conduct pre-attack exploration of vulnerable systems, employing techniques such as device finger-printing, port scanning [1, 12], social engineering [8], and similar methods.

Toward this end, we propose a Stackelberg game played among a set of interconnecting agents that include both defenders and attackers. The defenders are simultaneous first movers and the attackers are simultaneous second movers best responding to the defenders. The defenders' actions are both in anticipation of the attackers' best response and in anticipation of the impact from other defenders' actions through the interaction graph. This leads to the investigation of a special type of *Subgame Perfect Equilibrium (SPE)* as the solution concept, and is obtained through the backward induction which now involves solving a set of fixed point equations in both stages of the induction (by the defenders and the attackers).

It's worth noting that Stackelberg games as a formalism also enjoys very wide use in the cybersecurity literature. While the network game (or IDS game) is often used to study security investment and externality, the Stackelberg game has been used to model a variety of attacker-defender scenarios, see e.g., [4, 16]. To the best of our knowledge, this study is the first attempt at overlaying a Stackelberg game on top of a network game. Our main findings are as follows.

1. We first show that for a special class of interior SPE, the leaders' first stage game in the sequential-move network game is equivalent to a simultaneous-move network game played by the leaders on the Schur complement [32] of the interdependency (sub)matrix of the attackers in the full interdependency matrix (Theorem 4.1). We refer to this as the reduced network game, and use its properties to gain insights into leaders' behavior in the SPE of the sequential network game.

2. We then study a sequential network game with multiple defenders and a single attacker, and contrast the sequential-move game (i.e., defenders move first and the attacker moves second) with a simultaneous-move game (i.e., all players choose actions simultaneously). We first identify conditions under which the sequential nature of the game induces lower free-riding behavior and lower investment by the defenders (Propositions 5.1 and 5.2). This happens when the interdependency network of the defenders is a "rank-1" network (defined in Section 5.1) or if the influence of the attacker is significant. We show that the magnitude of this decrease aligns with the alpha-centrality of defenders in the reduced network (Proposition 5.3).

3. We further identify instances under which the opposite happens, i.e., when sequentiality leads some defenders to increase their investment, which may in turn lead the attacker to also increase its attack effort. Specifically, we illustrate a situation where severe free-riders (those substantially "supported" by other defenders' investments in the simultaneous-move game) are "exposed", with a diminished ability to free-ride in the sequential setting (Section 5.2). Intuitively, such defenders invest more in the SPE than the corresponding NE due to losing support from other defenders who invest less in the sequential-move game. We then show that as a result of this increase in effort by exposed free-riders, the attacker may best-respond by increasing its attack effort (Section 5.4).

The remainder of the paper is organized as follows. We review the literature most closely related to this work in Section 2. We present the network-Stackelberg game model in Section 3, followed by its equilibrium analysis in Section 4. We then contrast sequential vs. simultaneous-move multi-defender single-attacker games in Section 5. We discuss the limitations of our model and directions of future works in Section 6, and conclude the paper in Section 7.

2 Related Works

Sequential-move games, or Stackelberg games, have been widely studied in the economics literature. The most closely related to this work are those that have studied Stackelberg public good provision or network games, exploring the benefits to the leaders and/or contrasting the outcomes with those of simultaneous-move games. Sherali [27] studied a sequential game in the oligopoly market with identical firms who differ only in the sequence of participation. Sherali observed that leaders always make higher profit than followers, and each leader's profit is maximized when it is the only leader of the game. These results shed light on the leaders' sequential exploitation of the followers, but focus only on competitive relationship between agents. Varian [30] constructed a quasilinear model on 2-agent sequential public-good games and examined subsidy mechanisms that can yield Lindahl allocations. Varian concluded that without mechanisms, the provision of public good and the sum of utilities would be less than those in the simultaneous counterpart. However, opposite results are observed in several papers in experimental economics on sequential step-level public good games, such as Erev and Rapoport [10] and Normann [23]. Both experiments show that sequential step-level games actually improve public-good provision, even though second movers often punish first-movers for insufficient investment. The reason for the improvement, as suggested by Normann [23], is the *role model* effect: when the first mover tends to invest more in the public good, it serves as a role model for the second mover, who in turn would invest more. These different observations on the impact of sequential setting on the provision of public goods are to an extent due to the difference of the models, where [30] considers continuous and quasilinear utilities, while those in [10] and [23] are discrete. Additionally, these models only account for positive externalities between players and do not adequately incorporate a combination of positive and negative externalities, which can be effectively modeled through the framework of network games [7, 20, 22, 25]. To our knowledge, our work is the first to analyze network game played in sequential order, and contrast them with their simultaneous-move counterparts.

3 Model Description

We consider a public good provision game on network $\mathcal{G}$ played in a sequence of two sets of moves, the first by a set of leaders $\mathcal{L}$ and the second by a set of followers $\mathcal{F}$. Each player or agent i chooses an action $x_i \in \mathbb{R}_+$ denoting its level of investment/effort in the public good. All players are interconnected by the *game matrix* G, the $(i,j)^{th}$ entry of which denotes the dependence of player i's utility on player j's action. Thus, it is sometimes called a *dependence matrix* and conversely, G^T is sometimes called an *influence matrix*. Accordingly, whenever there is an edge pointing from i to j, we say player i *depends* on player j, or player j *influences* player i. Without loss of generality, we assume all diagonals of G are 1, i.e., $g_{ii} = 1, \forall i$. The matrix G is equivalent to a directed graph with edge weights representing both the connectivity among agents and the strengths of their connections. In graph theory terminology, the matrix $G - I$ is the adjacency matrix of the network with I being the identity matrix.

Since the agents separate into leaders and followers, it makes sense to divide the matrix G in blocks as follows:

$$G = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \,, \tag{1}$$

where A and D represent the dependency within each player set $\mathcal{L}$ and $\mathcal{F}$, respectively, while B and

C represent the (cross-)dependency between player sets $\mathcal{L}$ and $\mathcal{F}$. For the most part, we will use notations like a_{ij} and b_{ij} so as to make it clear the types of agents the edge weight is connecting; occasionally we will use g_{ij} to refer to a generic element in the full matrix.

Let $\boldsymbol{x}_l$ be the action profile (a column vector) of all leaders and $\boldsymbol{x}_f$ be the action profile of all followers. Denote by $\boldsymbol{x} := (\boldsymbol{x}_l^T, \boldsymbol{x}_f^T)^T$ the joint action profile of all players. In this two-stage setting, the leaders can be viewed as solving an equilibrium problem in a standard network game played on the network A, anticipating the followers' reaction to their own choices, i.e., $\boldsymbol{x}_f(\boldsymbol{x}_l)$, while the followers can be seen as solving an equilibrium problem in another standard network game played on D but with $\boldsymbol{x}_l$ as fixed parameters. More precisely, we adopt the following players' utility functions:

$$u_i(x_i, \boldsymbol{x}_{-i}|G) = f_i\left(x_i + \sum_{j\in\mathcal{L}\cup\mathcal{F}\setminus\{i\}} g_{ij}x_j\right) - \beta_i x_i, \quad \forall i \in \mathcal{L}\cup\mathcal{F}, \tag{2}$$

where $f_i(\cdot)$ is the benefit function of player i and $\beta_i > 0$ is player i's unit cost of effort or investment. We will assume $f_i(\cdot)$ is continuously differentiable, strictly increasing, and strictly concave for all i. We also assume that $f_i'(\cdot)$ is a convex function. This type of utility is also commonly known as the total weighted effort model in the literature of public good games, see e.g., Grossklags et al. [11] and Varian [29], where the argument of the utility function $f_i(\sum_j g_{ij}x_j)$ is called the *total weighted effort* received by player i.

Before proceeding, we introduce the following assumption to make our problem more tractable.

Assumption 3.1. *For every player $i \in \mathcal{L}\cup\mathcal{F}$, $f_i'(-\infty) = \infty$ and $f_i'(\infty) = 0$. That is, there always exists $q_i \in \mathbb{R}$ such that $f_i'(q_i) = \beta_i$. We will denote $\boldsymbol{q}_l := (q_i)_{i\in\mathcal{L}}$ and $\boldsymbol{q}_f := (q_i)_{i\in\mathcal{F}}$, and assume $\boldsymbol{q}_l, \boldsymbol{q}_f \geq 0$ for simplicity.*

Assumption 3.1 states that every player would be satisfied once its received total weighted effort reaches q_i. Since q_i equalizes i's marginal benefit and marginal cost, any additional investment yielding a total weighted effort higher than q_i is strictly less preferred by player i. Thus, we say q_i is the *saturation point* of player i.

A few comments are in order. First, in this model each agent has a scalar action, interchangeably termed the *effort* or *investment*. If we view agents in $\mathcal{F}$ as attackers and those in $\mathcal{L}$ as defenders, then this scalar action should be interpreted as a generic, non-targeted effort: common broad-based security mechanisms for a defender (such as firewalls, intrusion detection systems, etc.), and for an attacker, investment in attack technology which is then applied to all defenders. How susceptible a specific defender is to a specific attacker (the edge weights hard-coded in the game parameters) encapsulates factors not explicitly modeled; e.g., an entity with a large number of employees may be particularly susceptible to an attacker specializing in phishing, but less so to an attacker specializing in exploiting unpatched system vulnerabilities.

Secondly, there are two sets of terminologies frequently used in such a network setting that deserve some clarification, as we will repeatedly use these later on: *positive* vs. *negative externality*, and whether one's effort acts as a *strategic complement* or *strategic substitute* to that of another. In general, an action by a player is said to carry positive externality if the same, self-serving action also benefits others that depend on it. In the context of a game defined on a graph, this is often associated with a positive edge weight on an incoming edge to the player taking the action, where an increase in effort by this player benefits its neighbor who depends on this player; this is clearly seen in the total weighted effort model given above. Conversely, negative externality is typically associated with a negative edge weight. Whether the effort being provisioned is a substitute or complement has to do with whether the increase in a player's effort would lead a dependent neighboring player to decrease or increase its own effort, substitute in the former and complement in the latter. This, too, is closely associated with the signs of the edge weights: often a positive edge weight means substitute and negative means complement. However, the interpretation of both sets of concepts and their association with the edge weights ultimately depends on the utility function one adopts. Suffice to say that the total weighted effort model presented above does allow us to associate positive (resp. negative) edge weights with both positive (resp. negative) externality and actions being strategic substitutes (resp. complements) as described above.

Figure 1 depicts the relationship between these concepts and the saturation point in a 2-player example, where a single edge points from player 1 to player 2 with an edge weight of e. This is shown through the *best response* (BR) of player 1 to player 2's action x_2, which is taken as given in this example. In total weighted effort models with linear costs, this best response function is linear and given by $BR_1(x_2) = \max\{0, q_1 - ex_2\}$. This is shown in two cases: when $e > 0$ (the red curve)

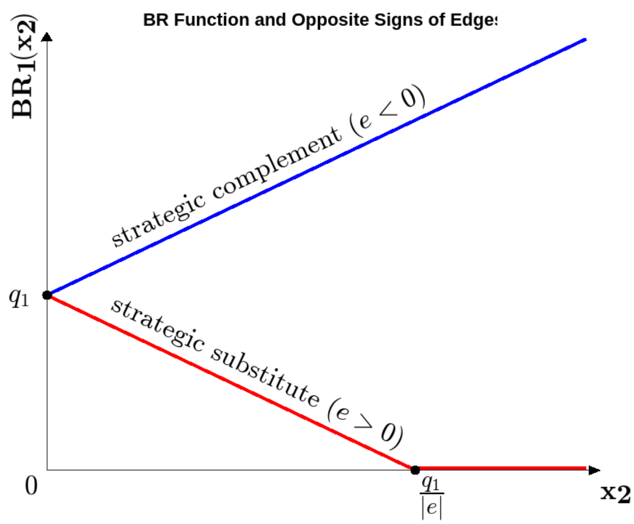


Figure 1: Relationship between saturation point and strategic substitute/complement.

and when $e < 0$ (the blue curve). Specifically, when $e > 0$, the amount that 1 has to invest to reach its saturation point is reduced by $|e|x_2$, the share in player 1's total weighted effort that comes from player 2. Therefore, player 1 would always select an effort level less than q_1, free-riding on player 2. In the extreme case, when $x_2 \geq \frac{q_1}{|e|}$ (the flat segment of the red curve), then player 1's total weighted effort has already reached its saturation point q_1 and hence has no incentive to take any positive action. This is a case of player 2's action x_2 *substituting* for 1's own action. In the opposite case, when $e < 0$, any positive x_2 lowers player 1's total weighted effort, forcing it to invest an additional

$|e|x_2$ to compensate for this loss, resulting in a x_1 higher than q_1, forming a *strategic complement* relationship.

In the context of cybersecurity, the total weighted effort model captures the fact that one player's security investment generally contributes to another, connected player's security posture. Sometimes this contribution is positive, e.g., a more secure service provider enhances the security of a customer that depends on its service. Other times this can be negative, e.g., strong defense by one player could induce threats to be directed at other players. In our analysis, we will generally take the view that security investment has positive externality among defenders, whereas attack investment carries negative externality for defenders.

Our goal with the two-stage network game model laid out in Eq. (2) is to understand how the resulting equilibrium compares to that of a typical, simultaneous-move public good game, to which our model is a natural extension. In a standard total weighted effort game, the Nash equilibrium is obtained by finding the fixed point of all agents' linear best response functions. It is worth noting that this step is much more complicated in the proposed sequential network game. This is because the best response mappings for the leaders are no longer linear, which will be discussed at length in Section 4.

4 Equilibrium Analysis

The problem outlined in the previous section is a two-stage dynamic (Stackelberg) game with perfect information, where the standard solution concept is the sub-game perfect equilibrium (SPE). An action profile $\boldsymbol{x}^*$ is a sub-game perfect equilibrium if every sub-game with $\boldsymbol{x}^*$ constrained to that sub-game forms a Nash equilibrium (NE). Sub-game perfect equilibrium in Stackelberg games are commonly solved using backward induction. We will limit our attention to an (easier) sub-class of sub-game perfect equilibrium and utilize the backward-induction technique to resolve it.

Given any fixed leaders' profile $\boldsymbol{x}_l$, the followers' game becomes a standard (simultaneous-move) network game. Naghizadeh and Liu [22] showed that due to strict concavity of utility functions in the second stage, $\boldsymbol{x}_f^*(\boldsymbol{x}_l)$ is a Nash equilibrium among the followers if and only if the following system of equations hold simultaneously

$$x_i^*(\boldsymbol{x}_l) = \max\left\{0,\ q_i - \sum_{j\in\mathcal{L}} c_{ij}x_j - \sum_{\substack{i'\in\mathcal{F}\\ i'\neq i}} d_{ii'}x_{i'}^*(\boldsymbol{x}_l)\right\}, \quad \forall i\in\mathcal{F}. \tag{3}$$

Notice Eq. (3) might not be the same for every leaders' action profile $\boldsymbol{x}_l$. Thus, we shall call $\boldsymbol{x}_f^*(\boldsymbol{x}_l)$ the followers' *equilibrium reaction function* to $\boldsymbol{x}_l$. [22] provides a necessary and sufficient condition for the existence and uniqueness of the solution to Eq. (3); we state this as an assumption below, which allows us to have a unique followers' equilibrium response, leading to a well defined sub-game perfect equilibrium introduced shortly.

Assumption 4.1. *The matrix D is a P-matrix.*

Assumption 4.1 ensures that the followers' equilibrium response is unique for any $\boldsymbol{x}_l$. Without this assumption, if the followers have multiple equilibrium responses, without a clearly defined tie breaking rule a leader would not know which response to in turn respond to and which response the other leaders would respond to. To avoid this complication, we will adopt Assumption 4.1 as a global assumption throughout the paper.

Following the technique of backward induction, we substitute Eq. (3) into the leaders' utility functions in Eq. (2) and obtain a simultaneous-move game among the leaders, which we refer to as the *reduced game* formally defined below.

Definition 4.1. *The reduced game is the simultaneous-move game among the leaders, obtained after one step of backward induction, i.e., replacing the terms $\{x_j : j \in \mathcal{F}\}$ in their utility functions (2) with the followers' equilibrium reaction functions $\{x_j^*(\boldsymbol{x}_l) : j \in \mathcal{F}\}$.*

By the standard definition, any Nash equilibrium of the reduced game is a sub-game perfect equilibrium. However, the fact that the followers can choose boundary actions, i.e., $x_j = 0$ for some $j \in \mathcal{F}$, complicates the application of standard theorems on the existence and computation of Nash equilibrium. In the following paragraphs, we briefly explain this difficulty and introduce a more restrictive equilibrium concept to sidestep this problem.

We begin by noting that the *max* operator in Eq. (3) makes this function non-differentiable, which leads to a fundamental difficulty in our model. When each leader i performs backward-induction by plugging in the followers' reaction function, its utility function (2) can be non-differentiable and non-concave with respect to its own investment x_i, both of which are important assumptions for most theorems on existence and uniqueness of Nash equilibrium, such as fixed-point theorems and variational inequalities. These results rely on the first-order condition reformulation of every player's optimization problem, which no longer holds when the utility functions are non-concave. Moreover, non-concavity also obstructs the construction of sub-differentials which are commonly used in solving non-smooth Nash equilibrium problems. We illustrate this difficulty through the example below.

Example 4.1. *Assume $\mathcal{L} := \{1\}$ and $\mathcal{F} := \{2\}$ with $q_2 = 1$[1] and $G = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$ for some $a \in (0, 1)$. Also assume $f_1(\cdot) = \log(\cdot)$. Thus, the follower's equilibrium response function is $x_2^*(x_1) = \max\{0, 1 - ax_1\}$. Using backward induction to derive the first-stage utility we obtain $u_1(x_1|G) = \log\left(x_1 + a \max\{0, 1 - ax_1\}\right) - \beta_1 x_1$. This is clearly non-differentiable at $x_1 = 1/a$. Since u_1 is at least continuous, we can assume it is right-differentiable and calculate its right-derivative, denoted by u_1^+, as*

$$u_1^+(x_1|G) = \frac{1 - a^2 \mathbb{1}_{\{x_1 \leq 1/a\}}}{x_1 + a \max\{0, 1 - ax_1\}} - \beta_1. \tag{4}$$

[1] We have started directly from q_2, skipping the formulation of $f_2(\cdot)$ and β_2. The subsequent analysis holds for any $f_2(\cdot)$ and β_2 that lead to $q_2 = 1$.

u_1 is concave if and only if u_1^+ is weakly decreasing. However, if we take arbitrarily small $\varepsilon > 0$, then

$$\frac{u_1^+((1+\varepsilon)/a|G) + \beta_1}{u_1^+((1-\varepsilon)/a|G) + \beta_1} = \frac{a/(1+\varepsilon)}{a(1-a^2)/[1-\varepsilon(1-a^2)]} = \frac{(1-a^2)^{-1} - \varepsilon}{1+\varepsilon} > 1, \tag{5}$$

which implies $u_1^+((1+\varepsilon)/a|G) > u_1^+((1-\varepsilon)/a|G)$. Thus, u_1 is not concave at the non-differentiable point $x_1 = 1/a$. In fact, when the leader chooses $x_1 = 1/a$, the follower's reaction becomes a boundary solution, i.e., $x_2^*(1/a) = 0$. In that case, the leader's utility function is concave and differentiable within the regions $[0, 1/a]$ and $(1/a, \infty)$ instead of on the entire domain. We will show that this issue can make the leader's best-response mapping non-convex, which hinders the application of, say, Kakutani's fixed-point theorem.

Let $a \in (1/\sqrt{2}, \sqrt{1-e^{-1}})$ and $\beta_1 = \frac{1-a^2}{a}\log\left(\frac{1}{1-a^2}\right)$. We apply the first-order (necessary) condition to the leader's problem and obtain two local solutions: $x_1^1 = \frac{1}{\beta_1}$ and $x_1^2 = \frac{1}{\beta_1} - \frac{a}{1-a^2}$. It can be verified that $x_1^1 \in (1/a, \infty)$ and $x_1^2 \in [0, 1/a]$. When the leader chooses x_1^1, the follower is inactive with $x_2^*(x_1^1) = 0$, but when the leader chooses x_1^2, the follower will respond with $x_2^*(x_1^2) = \frac{1}{1-a^2}\left(1 + \frac{a^2}{\log(1-a^2)}\right) > 0$. Nevertheless, in both cases the leader obtains the same utility since

$$u_1(x_1^2|G) = \log\left(\frac{1-a^2}{\beta_1}\right) - 1 + \frac{a\beta_1}{1-a^2} = \log\left(\frac{1}{\beta_1}\right) - 1 = u_1(x_1^1|G). \tag{6}$$

Thus, the solution set to the leader's optimization problem consists of two distinct points x_1^1, x_1^2. If this situation occurs for some player in a multi-leader game, then we cannot apply, e.g., Kakutani's fixed-point theorem, to identify a Nash equilibrium since its best-response function does not necessarily map to a convex set.

In light of the difficulty discussed above, we will focus on a sub-class of sub-game perfect equilibrium where all players choose interior actions, to which we refer as an *interior sub-game perfect equilibrium* defined below.

Definition 4.2. *An action profile $\boldsymbol{x}^*$ is called an interior sub-game perfect equilibrium (I-SPE) if it is sub-game perfect with $\frac{\partial u_i}{\partial x_i}(\boldsymbol{x}^*|G) = 0$ for every $i \in \mathcal{L} \cup \mathcal{F}$.*

For reference, an *interior Nash equilibrium* for a standard multi-player simultaneous-move game with utility functions given by $u_i(\boldsymbol{x})$ for every player i is defined as follows.

Definition 4.3. *An action profile $\boldsymbol{x}^*$ is called an interior Nash equilibrium (I-NE) if it is a Nash equilibrium and $\frac{\partial u_i}{\partial x_i}(\boldsymbol{x}^*) = 0$ for every player i.*

Note that, in Definition 4.2, we require the sub-game perfectness since the vanishing first-order condition is not equivalent to global optimality of each leader's best response. To establish the existence and uniqueness of I-SPE, we need to rule out cases where some followers exert zero effort (also called inactive). Below we introduce a sufficient condition for the existence and uniqueness of an I-SPE.

First define $G/D := A - BD^{-1}C$ as the Schur complement [32] of the block D in G. For each $i \in \mathcal{L}$, if $(G/D)_{ii} > 0$, then define $\tilde{q}_i$ such that $f_i'(\tilde{q}_i) = \frac{\beta_i}{(G/D)_{ii}}$. This is the saturation point of leader

i in the reduced game. If every $\tilde{q}_i$ exists, we stack them into a column vector $\tilde{\boldsymbol{q}}_l := (\tilde{q}_i)_{i\in\mathcal{L}}$ and will call this the *reduced saturation point*: it is the saturation point of the reduced game, and as we will see later it is also smaller in quantity.

Theorem 4.1. *Let G be a P-matrix with $C \leq 0$ and D being a Z-matrix, i.e., $d_{ij} \leq 0$ for all $i \neq j$. Then, $\tilde{\boldsymbol{q}}_l$ is well defined and there exists a unique SPE. Also, this SPE is also an I-SPE given by $\boldsymbol{x}^* = G^{-1}\begin{bmatrix}\tilde{\boldsymbol{q}}_l \\ \boldsymbol{q}_f\end{bmatrix}$, if $G^{-1}\begin{bmatrix}\tilde{\boldsymbol{q}}_l \\ \boldsymbol{q}_f\end{bmatrix} \geq 0$.*

Proof. **Step 1.** We first show that all followers are going to play interior actions. It is known that matrices that are both Z-matrices and P-matrices are nonsingular M-matrices [28]. This, together with Assumption 4.1, imply that D is a non-singular M-matrix. Furthermore, the inverse of a non-singular M-matrix is a non-negative matrix [28]. Therefore, due to the positivity of $\boldsymbol{q}_f$ (Assumption 3.1) and $-C$, we know $\boldsymbol{x}_f^*(\boldsymbol{x}_l) = D^{-1}(\boldsymbol{q}_f - C\boldsymbol{x}_l)$ is the unique solution to the system (3) for every $\boldsymbol{x}_l \geq 0$. It can be verified that this is always an interior solution, i.e., $\frac{\partial u_i}{\partial x_i}(\boldsymbol{x}_l, \boldsymbol{x}_f^*(x_l)|G) = 0$ for every $i \in \mathcal{F}$ and $\boldsymbol{x}_l \geq 0$.

Step 2. Next we verify the existence and uniqueness of a Nash equilibrium in the reduced game. Knowing the followers always play interior actions, the leaders' utility function in the reduced game can be expressed in closed-form as

$$\tilde{u}_i(\boldsymbol{x}_l|G) := u_i(\boldsymbol{x}_l, \boldsymbol{x}_f^*(\boldsymbol{x}_l)|G) = f_i\left(\left[(G/D)\,\boldsymbol{x}_l + BD^{-1}\boldsymbol{q}_f\right]_i\right) - \beta_i x_i, \quad \forall i \in \mathcal{L}. \tag{7}$$

Since G is a P-matrix, G/D is also a P-matrix [28]. Thus, $(G/D)_{ii} > 0$ for all $i \in \mathcal{L}$ since they are the 1×1 principal minors of G/D, and hence $\tilde{\boldsymbol{q}}_l$ is well-defined. Furthermore, by replacing $f_i(\cdot)$ with $\tilde{f}_i(\cdot) := f_i(\cdot - t_i)$ where $t_i = [BD^{-1}\boldsymbol{q}_f]_i$, one can recover a standard simultaneous-move network game played over the network G/D. Since G/D is a P-matrix, this game has a unique Nash equilibrium.

Step 3. Finally, we show that this SPE is interior. The Nash equilibrium of the reduced game is equivalent to the solution of the following best-response relationship:

$$x_i = \frac{1}{(G/D)_{ii}} \max\left\{0,\ \tilde{q}_i - [BD^{-1}\boldsymbol{q}_f]_i - \sum_{\substack{i'\in\mathcal{L}\\ i'\neq i}} (G/D)_{ii'} x_{i'}\right\}, \quad \forall i \in \mathcal{L}. \tag{8}$$

By the matrix inversion lemma, we have

$$\begin{aligned} G^{-1}\begin{bmatrix}\tilde{\boldsymbol{q}}_l \\ \boldsymbol{q}_f\end{bmatrix} &= \begin{bmatrix}(G/D)^{-1} & -(G/D)^{-1}BD^{-}1 \\ -D^{-1}C(G/D)^{-1} & D^{-1} + D^{-1}C(G/D)^{-1}BD^{-1}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{q}}_l \\ \boldsymbol{q}_f\end{bmatrix} \\ &= \begin{bmatrix}(G/D)^{-1}(\tilde{\boldsymbol{q}}_l - BD^{-1}\boldsymbol{q}_f) \\ D^{-1}\left[\boldsymbol{q}_f - C(G/D)^{-1}(\tilde{\boldsymbol{q}}_l - BD^{-1}\boldsymbol{q}_f)\right]\end{bmatrix} =: \begin{bmatrix}\boldsymbol{v}_1 \\ \boldsymbol{v}_2\end{bmatrix}. \end{aligned} \tag{9}$$

By the stated conditions, $\boldsymbol{v}_1, \boldsymbol{v}_2$ are both non-negative. One can easily verify that $\boldsymbol{v}_1$ is the solution to Eq. (8). Thus, it is the unique Nash equilibrium of the reduced game. Furthermore, $\boldsymbol{v}_2$ is exactly the equilibrium reaction of the followers to the leaders' action profile $\boldsymbol{v}_1$. Thus, $(\boldsymbol{v}_1^T, \boldsymbol{v}_2^T)^T$ constitutes the unique SPE. Besides, since $\frac{\partial u_i}{\partial x_i}(\boldsymbol{v}_i|G) = 0$ for $i \in \mathcal{L}$ and $\frac{\partial u_j}{\partial x_j}(\boldsymbol{v}_2) = 0$ for $j \in \mathcal{F}$, it is also the unique I-SPE of this game. □

Theorem 4.1 reveals an interesting connection between the sequential-move game and the simultaneous-move game on a network. When the followers are known to play interior actions, the leaders' reduced game is equivalent to a simultaneous-move game with game matrix G/D (the Schur complement of D in G). With a slight abuse of terminology, we will also refer to the latter as the *reduced game* and G/D as the *reduced game matrix* for the rest of the paper. Note the dimensionality reduction in the game matrix from the original game, i.e., $dim(G/D) = dim(A) < dim(G)$.

5 Networked Security: A Multi-Defender Single-Attacker Game

We now turn our attention to applying this sequential network game framework to the application of networked security (investment). To do so, we will view the leaders as simultaneous defenders, who set their defense strategies $\boldsymbol{x}_l$ beforehand and in anticipation of attacks, and view the followers as simultaneous attackers who choose optimal attack efforts $\boldsymbol{x}_f$ in response. In addition, we will assume that the edge weights between any defender and any attacker are negative in both directions, if they exist and that the conditions in Theorem 4.1 are satisfied. Formally,

Assumption 5.1. *The following conditions hold:*

- *B and C are non-positive matrices,*
- *D is a Z-matrix,*
- *G is a P-matrix,*
- $G^{-1}\begin{bmatrix}\tilde{\boldsymbol{q}}_l \\ \boldsymbol{q}_f\end{bmatrix} \geq 0$.

A direct consequence of Assumption 5.1 is the decrease in the defenders' saturation points. We illustrate this in Figure 2 and provide a formal proof in the following lemma.

Lemma 5.1. $\tilde{\boldsymbol{q}}_l \leq \boldsymbol{q}_l$.

Proof. According to the proof of Theorem 4.1, D^{-1} is a non-negative matrix. Thus, the matrix $BD^{-1}C$ is also non-negative. Then, $(G/D)_{ii} = 1 - [BD^{-1}C]_{ii} \leq 1$. Thus, for every $i \in \mathcal{L}$, $\frac{\beta_i}{(G/D)_{ii}} \geq \beta_i$. At the same time, the concavity of f_i implies that $f_i'(\cdot)$ is decreasing. This combined with the definition $f_i'(q_i) = \beta_i$ and $f_i'(\tilde{q}_i) = \frac{\beta_i}{(G/D)_{ii}}$ means $\tilde{q}_i \leq q_i$. □

We now take an in-depth look at this game in the special case of a single attacker. The network matrix in this case can be written as $G = \begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{c}^T & 1 \end{bmatrix}$, where $\boldsymbol{b}$ and $\boldsymbol{c}$ are non-positive column vectors. Then, by Theorem 4.1, the simultaneous-move game between the defenders, obtained after one step

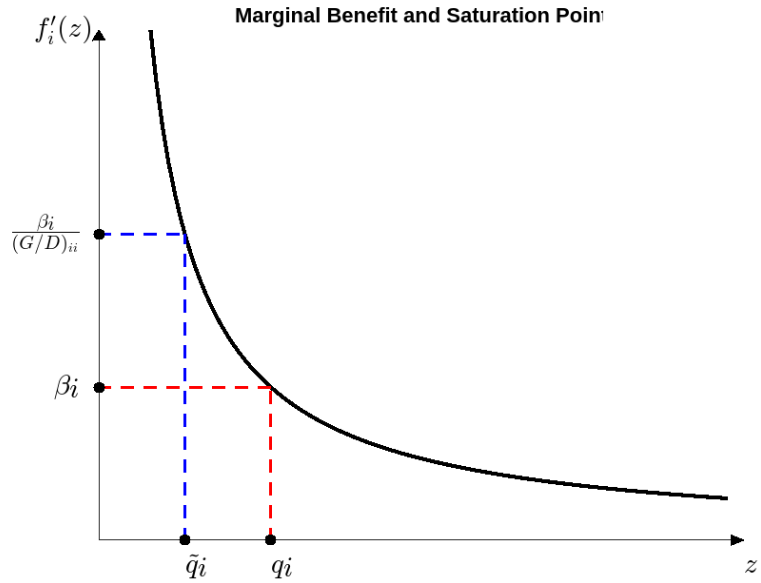


Figure 2: The drop in the saturation point in the sequential setting.

of backward induction, can be viewed as the reduced game over the reduced game matrix $A - \boldsymbol{b}\boldsymbol{c}^T$ (the Schur complement of D in G with $D = 1$). Under Assumption 5.1, the edge weights in this Schur complement are less than A. Thus, we immediately see a *strategic complement effect* induced by the presence of the follower/attacker. Interestingly, if some edges are positive in the original game but become negative in the reduced game, then it means that *the presence of the attacker has turned strategic substitutes into strategic complements*, or in other words, allies into adversaries.

In the remainder of this section, we investigate how the sequentiality affects players' equilibrium actions and free-riding behaviors under different network topologies, including the strategic complement effect mentioned above. Specifically, Subsection 5.1 examines general conditions under which the defenders' free-riding and equilibrium efforts are reduced. Subsection 5.2 shows an opposite instance where sequentiality exposes free-riders by inducing them to invest more. Then, Subsection 5.3 studies the connection between equilibrium effort changes and the alpha-centrality of defenders in the network. Finally, Subsection 5.4 examines the impact of the free-riding behavior among defenders on the attacker's equilibrium response.

We begin by calculating the unique I-SPE of the sequential game using Theorem 4.1, which is given by

$$\boldsymbol{x}_l^{SPE} = (A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}(\tilde{\boldsymbol{q}}_l - q_f\boldsymbol{b}) \ , \tag{10}$$

$$x_f^{SPE} = \left(1 + \boldsymbol{c}^T(A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\boldsymbol{b}\right) q_f - \boldsymbol{c}^T(A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\tilde{\boldsymbol{q}}_l. \tag{11}$$

Suppose, on the other hand, that all players (all the defenders and the attacker) play simultaneously with the same set of game parameters and that $G^{-1}\begin{bmatrix}\boldsymbol{q}_l \\ q_f\end{bmatrix} \geq 0$. Then, there exists a unique Nash equilibrium, which is also interior, given by Eqs. (10) and (11), with $\tilde{\boldsymbol{q}}_l$ replaced by $\boldsymbol{q}_l$. We will denote these respectively as $\boldsymbol{x}_l^{NE}$ and x_f^{NE}.

5.1 The Interplay Between Sequentiality and Dependence

In this section, we examine how the presence of a strategic second-stage attacker affects the free-riding patterns among defenders. Denote by $\boldsymbol{x}_l^{SPE}(A)$ (resp. $\boldsymbol{x}_l^{NE}(A)$) the I-SPE (resp. I-NE) given the defenders' game matrix A. Define the *vector of free-riding scores* under both I-SPE and I-NE, respectively, as

$$\boldsymbol{s}^{SPE} = \boldsymbol{x}_l^{SPE}(I) - \boldsymbol{x}_l^{SPE}(A), \tag{12}$$

$$\boldsymbol{s}^{NE} = \boldsymbol{x}_l^{NE}(I) - \boldsymbol{x}_l^{NE}(A), \tag{13}$$

where I denotes the identity matrix. The i^{th} entry of $\boldsymbol{s}$ represents the decrease in investment of defender i in the interconnected network game, compared to that of an independent (non-networked) game. Therefore, the higher the value of s_i, the higher the degree of free-riding of defender i.

Let Assumption 5.1 hold for A replaced by I. Then, in light of Eq. (10), these score vectors can be expressed as

$$\boldsymbol{s}^{SPE} = \left((I - \boldsymbol{b}\boldsymbol{c}^T)^{-1} - (A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\right)(\tilde{\boldsymbol{q}}_l - q_f\boldsymbol{b}), \tag{14}$$

$$\boldsymbol{s}^{NE} = \left((I - \boldsymbol{b}\boldsymbol{c}^T)^{-1} - (A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\right)(\boldsymbol{q}_l - q_f\boldsymbol{b}). \tag{15}$$

We are interested in the difference $\Delta\boldsymbol{s} := \boldsymbol{s}^{SPE} - \boldsymbol{s}^{NE}$, i.e., the increment in the degree of free-riding induced by sequentiality. If the i^{th} entry of $\Delta\boldsymbol{s}$ is positive, then the free-riding behavior is amplified in the sequential version; otherwise, it is reduced. We will only focus on games of strategic substitutes among the defenders since free-riding behaviors are well-defined in these situations.

Proposition 5.1. *If A consists of purely strategic substitutes (i.e., $a_{ij} \geq 0, \forall i, j \neq i$) with small magnitude compared to the attacker's influence ($a_{ij} \leq b_i c_j, \forall i, j \neq i$), then there is less free-riding for all defenders in the sequential game, i.e., $\boldsymbol{s}^{SPE} \leq \boldsymbol{s}^{NE}$. Moreover, all defenders will exert lower effort in the sequential-move game compared to the simultaneous-move game, i.e., $\boldsymbol{x}_l^{SPE} \leq \boldsymbol{x}_l^{NE}$.*

Proof. It is sufficient to show that $\Delta\boldsymbol{s} = \left((I - \boldsymbol{b}\boldsymbol{c}^T)^{-1} - (A - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\right)(\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l) \leq 0$. Denote $X := I - \boldsymbol{b}\boldsymbol{c}^T$ and $Y := A - \boldsymbol{b}\boldsymbol{c}^T$. Also, let $P := X^{-1}$ and $Q := Y^{-1}$. Assumption 5.1 and the condition above implies X and Y are both non-singular M-matrices. Thus, $P, Q \geq 0$ by [28]. Since $\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l \leq 0$, it is sufficient to show $P - Q$ is a non-negative matrix, or equivalently, $p_{ij} \geq q_{ij}$ for all i, j. Since $XP = I$, we have

$$\sum_{k=1}^{N} x_{ik} p_{ki} = 1 \quad \forall i, \quad \text{and} \quad \sum_{k=1}^{N} x_{ik} p_{kj} = 0 \quad \forall i \neq j. \tag{16}$$

Notice $x_{ii} = y_{ii}$ for all i and $x_{ij} \leq y_{ij}$ for all $i \neq j$. Thus,

$$\sum_{k=1}^{N} y_{ik} p_{ki} \geq 1 \quad \forall i, \quad \text{and} \quad \sum_{k=1}^{N} y_{ik} p_{kj} \geq 0 \quad \forall i \neq j. \tag{17}$$

Now consider entries of the matrix $P = QYP$:

$$p_{ij} = \sum_{k=1}^{N} q_{ik} \sum_{l=1}^{N} y_{kl} p_{lj} = q_{ij} \sum_{l=1}^{N} y_{jl} p_{lj} + \sum_{k \neq j} q_{ik} \sum_{l=1}^{N} y_{kl} q_{lj} \geq q_{ij}, \quad \forall i, j. \tag{18}$$

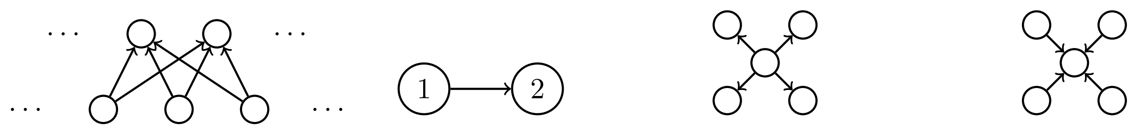
Therefore, $P - Q \geq 0$ and thus $\Delta \boldsymbol{s} = (P - Q)(\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l) \leq 0$. The last statement follows from $Q(\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l) \leq 0$. □

Proposition 5.1 indicates that the sequential nature of the game might induce lower investment by the defenders when they are able to exploit the follower to their own advantage, even though they are unable to free-ride on each other at the same time. The former is a common yet important observation in most sequential economic games, such as [27, 30], where the leaders' advantage arises as they no longer face the "threat of strategic uncertainty" posed by the followers. Consider, for instance, the single-defender and single-attacker case, whereby the defender's utility in an SPE is no lower than that in an NE as the former comes from a larger set that includes the latter. This tendency to decrease efforts is also consistent with the change in the defenders' saturation point, i.e., $\tilde{\boldsymbol{q}}_l \leq \boldsymbol{q}_l$ (Lemma 5.1). However, reasoning about this advantage in the general case is complicated by the presence of multiple defenders, because we now have to deal with the factor of free-riding which could lead to effort increase in the sequential game. To see this, consider one free-riding defender - the supported - who depends on the effort of another defender - the supporter. If the supporter's sequential effort drops according to the decrease in its saturation point, the supported receives less external efforts and now has to invest more by itself to compensate for it. Therefore, whether the equilibrium effort of any defender decreases depends on the balance between these two factors. The conditions in Proposition 5.1 limit the defenders' reliance on free-riding (on other defenders), so that the effect of lower saturation point dominates, leading to lower effort by all defenders in the sequential setting.

Furthermore, Proposition 5.1 shows an interesting phenomenon where the strategic substitute network with small dependencies induces lower investments in sequential games but also less free-riding. In fact, the reduced game consists of no positive dependencies, namely free-riding, at all! This is because the strategic complement effect turns off-diagonal values of the Schur complement of D in G from non-negative to non-positive. This is exactly the "turning-allies-into-adversaries" situation mentioned at the beginning of this section.

When substitute dependencies are large, there exist networks with special topologies that can similarly reduce free-riding scores for all defenders. We refer to these as "rank-1" networks, literally meaning that the adjacency matrix of such a network is of rank 1. Consider a strategic substitute network with its adjacency matrix given by $A := \boldsymbol{u}\boldsymbol{v}^T$ where $\boldsymbol{u}, \boldsymbol{v} \geq 0$. Then, as the diagonals of A are all zero, we derive the following complementarity condition: either $u_i = 0$ or $v_i = 0$, $\forall i$. The interpretation of $u_i = 0$ is the i-th agent depends on no one, while that of $v_i = 0$ is the i-th agent influences no one. Thus, "rank-1" networks possess the important "transient-node-free" property, where every node exclusively contains either outgoing or incoming edges, but not both. In other words, every node is either a source (node consisting only of outgoing edges) or a sink (node consisting only of incoming edges), in graph terminology.

The "transient-node-free" networks can be understood as a special case of directed bipartite graphs depicted in Figure 3a. Specifically, it is a directed graph with its vertex set divided into two disjoint subsets, each containing either source or sink nodes. In Figure 3a, the bottom nodes represent sources, while the top nodes represent sinks. Notice that this category includes some commonly seen topologies such as unidirectional, two-agent topology (Figure 3b) and star topology (Figure 3c and 3d). This category closely resembles a practical scenario where a group of agents relies on certain cyber-security services provided by another group of agents.



(a) Directed bipartite topology. **(b)** Two-agent topology. **(c)** Central source and star topology. **(d)** Central sink and star topology.

Figure 3: Examples of rank-1 networks.

Formally, the reduction of free-riding in rank-1 networks is shown in the following proposition.

Proposition 5.2. *Let the strategic substitute graph among defenders be of rank-1. Then free-riding is reduced among all defenders in the sequential case.*

Proof. As the defenders' graph has rank 1, by a generalization of Sherman-Morrison formula [2],

$$\begin{aligned}\Delta \boldsymbol{s} &= \left((I - \boldsymbol{b}\boldsymbol{c}^T)^{-1} - (I + \boldsymbol{x}\boldsymbol{y}^T - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\right)(\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l) \\ &= \frac{1}{1 + \boldsymbol{y}^T(I - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\boldsymbol{x}}(I - \boldsymbol{b}\boldsymbol{c}^T)^{-1}\boldsymbol{x}\boldsymbol{y}^T(I - \boldsymbol{b}\boldsymbol{c}^T)^{-1}(\tilde{\boldsymbol{q}}_l - \boldsymbol{q}_l).\end{aligned} \tag{19}$$

Note that $\boldsymbol{x}, \boldsymbol{y} \geq 0$, and $(I - \boldsymbol{b}\boldsymbol{c}^T)^{-1}$ is the inverse of a non-singular M-matrix, which is non-negative. Therefore, we have $\Delta \boldsymbol{s} \leq 0$. □

Proposition 5.2 states that even if the positive externalities among the defenders are non-negligible, i.e., the reduced network has some positive off-diagonal entries, as long as the defenders' network is of rank-1, all free-riding behaviors can be reduced in the sequential game. Consequently, this finding implies that in practical scenarios where the cyber-security of one group of agents relies on another (disjoint) group, the sequential version of the game tends to exhibit diminished free-riding behaviors, regardless of the strength of the dependency. Therefore, systems with such dependency structures, where the prevention of free-riding may be of major concern, can benefit from this type of sequentiality.

It is worth noting that a reduction in free-riding does not always coincide with a reduction in the defenders' efforts. This is demonstrated in the example discussed in the next subsection, where the game matrix has rank-1 (and so by Proposition 5.2, free-riding is lower in the sequential game) but one of the defenders invests more in the sequential game.

5.2 "Exposed" Free-Riders

Subsection 5.1 identifies the combined effect of free-riding and decrease in saturation points on the defenders' equilibrium effort, and in particular, conditions under which all defenders invest less and exhibit less free-riding. We now look at instances of the game where (some) defenders' efforts increase as a result of their diminished ability to free-ride on other defenders in a sequential game. We say the change from simultaneous-move to sequential-move *exposes* the free-riders; the following example illustrates this phenomenon.

Example 5.1. *In this example, we illustrate a situation where free-riders (those "supported" by other defenders' investments in the simultaneous-move game) are exposed by the sequential setting. We will see that such defenders invest more in the SPE than the corresponding NE due to losing support from other defenders who invest less in the sequential-move game.*

Let there be two defenders and one attacker. Given $a \in (0,1)$ and $b \in (-1,0)$, let the game matrix be such that $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 0 \\ b \end{bmatrix}$, $\boldsymbol{c} = \begin{bmatrix} b & b \end{bmatrix}$, and $d = 1$. The connection given in A indicates there is a one-way dependence between the two defenders: 1 depends on 2 but not the other way round. For simplicity, assume G is row diagonally dominant, which trivially implies the P-matrix property. The Schur complement of D in G for this game is given by $G/D = \begin{bmatrix} 1 & a \\ -b^2 & 1-b^2 \end{bmatrix}$. Let the benefit function be $f_i(\cdot) = \log(\cdot)$ for all $i = 1,2,3$. This choice guarantees that $q_i > 0$ for every player i. Then, we have $\tilde{q}_1 = q_1$, but $\tilde{q}_2 = (1-b^2)q_2$. The equilibrium efforts of the defenders for the SPE and NE are given by $\boldsymbol{x}_l^{SPE} = \det(G/D)^{-1} \begin{bmatrix} (1-b^2)q_1 - a(1-b^2)q_2 \\ b^2 q_1 + (1-b^2)q_2 \end{bmatrix}$ and $\boldsymbol{x}_l^{NE} = \det(G/D)^{-1} \begin{bmatrix} (1-b^2)q_1 - aq_2 \\ b^2 q_1 + q_2 \end{bmatrix}$, respectively. One can immediately check that defender 1 exerts higher effort in the sequential game while defender 2 exerts lower effort. Since defender 2 receives no positive externality, its equilibrium action is consistent with the decrease of its saturation point. On the other hand, defender 1, who receives positive externality from defender 2's effort, experiences a decrease of support from defender 2 in the sequential setting. What's worse, since defender 1 receives no negative externality from the attacker ($b_1 = 0$), its saturation point stays unchanged. Thus, it has to exert more effort in the sequential equilibrium.

5.3 Alpha-Centrality: Whose Effort Will Go Down?

We next present a method for identifying *which* defenders will tend to lower their investments in the sequential-move game. Eq. (10) and (11) suggest the following connection between such defenders and the alpha-centrality measure proposed by Bonacich and Lloyd [3]. Alpha-centrality is a centrality measure tailored for asymmetric matrices. Given parameters α, $\boldsymbol{v}$, and a graph with adjacency matrix G, the alpha-centrality is defined as

$$c_\alpha(G, \alpha, \boldsymbol{v}) = (I - \alpha G)^{-1}\boldsymbol{v}, \tag{20}$$

where α characterizes the strength of the transfer of "endogenous" status to vertices due to network connections with other vertices with high status, while $\boldsymbol{v}$ is interpreted as some "exogenous" status characteristic that can be different among vertices.

Proposition 5.3. *Consider the graph among defenders with the game matrix $A-\boldsymbol{b}\boldsymbol{c}^T$, with diagonals normalized to 1. Define the vector of exogenous status as $\Delta\boldsymbol{q}_l := \boldsymbol{q}_l - \tilde{\boldsymbol{q}}_l \geq 0$. Then the effort of defenders with higher alpha-centrality $c_\alpha(A-\boldsymbol{b}\boldsymbol{c}^T, -1, \Delta\boldsymbol{q}_l)$ decreases more in the sequential-move game compared to the simultaneous-move game.*

Proof. Subtracting $\boldsymbol{x}_l^{SPE}$ from $\boldsymbol{x}_l^{NE}$, we obtain $(A-\boldsymbol{b}\boldsymbol{c}^T)^{-1}\Delta\boldsymbol{q}_l$, which is exactly the alpha-centrality with $\alpha = -1$ and $\boldsymbol{v} = \Delta\boldsymbol{q}_l$. □

Note that the phrase "decreases more" only suggests the direction of change in equilibrium efforts; it does not exclude the possibility that defenders may increase their sequential investments, in which case the defender would have a negative alpha-centrality. A larger magnitude of such centrality would then indicate that the defender's effort increases more than others.

Proposition 5.3 provides several important insights on how defenders can exploit the sequential nature of the game. First, we note that an alpha-centrality with transfer strength $\alpha = -1$ induces an *alternating sign effect* as the influence of one node propagates along any path in the network [22]. For example, let $i_0, i_1, \ldots, i_k$ be a path of length k in the network, with the associated edge weights being $g_{i_0 i_1}, g_{i_1 i_2}, \ldots, g_{i_{k-1} i_k}$. Then, the effect of i_k's exogenous status on i_0's effort along this path is given by $(-1)^k \Delta q_{i_k} \prod_{l=1}^{k} g_{i_{l-1} i_l}$.[2] If the path length is even, the influence of i_k on i_0 along this path has the same sign as the product of the edge weights along the path, i.e., $\prod_{l=1}^{k} g_{i_{l-1} i_l}$; the effect is opposite otherwise. This observation suggests two fundamental properties of the network which characterize the influence between any two, possibly non-neighboring, agents: the product of the edge weights along the path connecting the agents, as well as the length of the path. The influence of one agent on the other is, therefore, the combined effect of all paths from the former to the latter; the exhibited investment reduction for every agent can be viewed as a combined effect of (besides the exogenous status $\Delta\boldsymbol{q}_l$) such influence of all other agents.

Those defenders who increase their investments receive negative alpha-centrality. Therefore, according to Subsection 5.2, these defenders are also free-riders in the simultaneous-move game.

Moreover, due to the convexity of $f_i'(\cdot)$, we have $q_i - \tilde{q}_i = (f_i')^{-1}(\beta_i) - (f_i')^{-1}\left(\frac{\beta_i}{(G/D)_{ii}}\right) \geq [(f_i')^{-1}]'\left(\frac{\beta_i}{(G/D)_{ii}}\right)\beta_i\left(1 - \frac{1}{(G/D)_{ii}}\right) = \frac{\beta_i}{f_i''(\tilde{q}_i)}\left(1 - \frac{1}{(G/D)_{ii}}\right)$. Thus, the curvature of the benefit function plays a role in each defender's exogenous centrality. The defenders with benefit functions of softer slopes (those that flatten-out more slowly) would receive a higher exogenous status characteristic, which might increase their alpha-centrality. This is because those defenders' saturation points are relatively farther away from zero compared to others. Thus, they can afford larger magnitude of investment drops whenever exploitation of sequentiality is possible.

[2] This expression can be intuitively understood by expanding the alpha-centrality $(I - \alpha G)^{-1}\boldsymbol{v}$ into $\sum_{k=0}^{\infty} \alpha^k G^k \boldsymbol{v}$ when $\alpha < \lambda_{\max}(G)^{-1}$. We refer interested readers to [3, 22].

5.4 Exposed Free-Riders Could Lead to Higher Attacker Effort

Note that for all multi-leader and single-follower security games, by comparing Eq. (10) and (11), we can see that $\boldsymbol{x}_l^{NE} - \boldsymbol{x}_l^{SPE} \geq 0$ implies $x_f^{NE} - x_f^{SPE} \geq 0$; that is, if all defenders' efforts drop in the sequential-move game (as happens, e.g., under the conditions of Proposition 5.1), then the attacker effort will drop as well. Conversely, this means if there is higher attacker effort in a sequential game instance, at least one defender must have increased its sequential investment in this game. By our insights from Subsection 5.2, a defender who increases its effort in a sequential-move game is likely to have been a free-rider in the simultaneous-move game. The following example elaborates on this observation about the relationship between the change in the defenders' free-riding behavior and the increases in the attacker's effort.

Example 5.2. *In this example, we present two instances of the security game where the attacker's sequential action increases in one and decreases in the other, compared to the simultaneous-move counterparts. We conclude from the example that the attacker will exert higher effort when there is more severe free-riding among the defenders.*

We begin by simplifying the last term in Eq. (11) into $\left(1 + \frac{\boldsymbol{c}^T A^{-1}\boldsymbol{b}}{1-\boldsymbol{c}^T A^{-1}\boldsymbol{b}}\right)(-\boldsymbol{c})^T A^{-1}\tilde{\boldsymbol{q}}_l$ *using the Sherman-Morrison formula. Thus,* $x_f^{NE} - x_f^{SPE} = \left(1 + \frac{\boldsymbol{c}^T A^{-1}\boldsymbol{b}}{1-\boldsymbol{c}^T A^{-1}\boldsymbol{b}}\right)(-\boldsymbol{c})^T A^{-1}\Delta\boldsymbol{q}_l$*, where* $\Delta\boldsymbol{q}_l$ *is the same vector of exogenous status in Proposition 5.3. Notice the term* $1 + \frac{\boldsymbol{c}^T A^{-1}\boldsymbol{b}}{1-\boldsymbol{c}^T A^{-1}\boldsymbol{b}}$ *is always positive, and therefore the change in the attacker's effort only depends on the value of* $(-\boldsymbol{c})^T A^{-1}\Delta\boldsymbol{q}_l$.

If the defenders are independent, i.e., $A = I$*, then* $(-\boldsymbol{c})^T A^{-1}\Delta\boldsymbol{q}_l = (-\boldsymbol{c})^T\Delta\boldsymbol{q}_l \geq 0$*. Thus, the attacker would exert less effort in the sequential game. If, on the contrary, we choose the network matrix to have one-way dependency, i.e.,* $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$*, the the attacker exerts strictly higher effort in the sequential-move game if* $(-\boldsymbol{c})^T A^{-1}\Delta\boldsymbol{q}_l = -c_1\Delta q_1 + (-c_2 + ac_1)\Delta q_2 < 0$*. This does not hold if either* $c_1 = 0$ *or* $\Delta q_2 = 0$*, where the latter is equivalent to* $b_2 = 0$ *or* $c_2 = 0$ *since* $\tilde{q}_i$ *is defined by* $f_i'(\tilde{q}_i) = \frac{\beta_i}{1-b_2c_2}$ *and* $f_i'(\cdot)$ *is strictly decreasing. Conversely, it requires* $\boldsymbol{c} < 0$ *and* $b_2 < 0$ *for the attacker to exert strictly higher effort in the sequential game. Moreover, if these conditions hold, we can rewrite the inequality into* $a > \frac{\Delta q_1}{\Delta q_2} + \frac{c_2}{c_1}$*. Notice that* $\frac{\Delta q_1}{\Delta q_2} + \frac{c_2}{c_1}$ *is strictly positive. Thus, only positive dependency among defenders can induce higher attacker effort, and the more positive this dependency is, the more effort the attacker would exert. Recall that, as we have discussed in Proposition 5.3, a higher value of* a *may lead to higher security investment of defender* 1*; this is the defender that free-rides on defender* 2 *in the simultaneous-move game, but is exposed in the sequential setting. Comparing these two findings, we conclude that the attacker is led to increase its effort in the sequential setting in response to the increase in effort of the "exposed" free-riding defender 1.*

6 Discussions

Our model provides a general framework of analyzing sequential-move network games and allows us to draw a number of interesting insights especially in a security game context. There are limitations to this model, which suggest possible directions of future research.

First of all, under the present model the attacker exerts a single (attack) effort which is felt by all (connected) defenders, modulo the connection strength. Since the connection strength is taken as a fixed game parameter, this model does not adequately capture the cases where attacks on individual defenders are highly targeted, often as a function of the defender's effort. This would require a different model. One possibility is to allow the attacker to choose optimal edge weights $\boldsymbol{b}$ and $\boldsymbol{c}$ as part of its strategic decisions.

Secondly, all our analysis is based on the assumption that an interior SPE exists. This means that our results do not directly apply to cases where some players may exert zero effort (inactive) or other boundary actions not captured by interior solutions. The latter is particularly important as all effort levels must ultimately be bounded due to budget constraints. We conjecture that, as long as the attacker remains interior, most of our insights should still hold because if we eliminate rows and columns of G corresponding to inactive defenders, we obtain exactly the same problem with a smaller graph. The challenge arises when there are inactive attackers, e.g., when the condition on D in Theorem 4.1 does not hold. In this case, the defenders' utility functions in the reduced game may not be concave, as we demonstrated in Example 4.1, and our results in Section 5 may not generalize.

One possible avenue is to formulate these problems as *equilibrium problems with equilibrium constraints* (EPEC) [31] and solve them numerically using algorithms for quasi-variational inequality, see e.g., [18, 24], or iterated best-response dynamics with diagonalization, see e.g. [31]. A more comprehensive investigation of these methods is a direction of future work.

Last but not least, it is particularly interesting to investigate sequential games with multiple defenders and multiple attackers. It is proposed that cyber criminals act economically in a dark market [17], meaning the number of attackers could be unbounded and their behaviors might be captured by traditional oligopoly models. We are interested in how the knowledge of multiple attackers may influence the security investment and free-riding behaviors among defenders. The technical challenge here is significant: under the proposed sequential network game model, the possibility of positive off-diagonal entries of D violates the conditions in Theorem 4.1, and thus it may require an alternative model.

7 Conclusion

In this paper we examined a sequential network game model that takes into account interdependency as well as sequential moves in the agents' strategy reasoning. Similar to general sequential public good games, this model encounters difficulties in the non-convexity and non-differentiability of the leaders' utility functions in the reduced games. We provided a sufficient condition under which an interior sub-game perfect equilibrium can be solved analytically.

In applying this model to the cyber-security context with a single attacker and multiple defenders, we demonstrated that the presence of a second-mover, the strategic attacker, can have strategic complement effects on the defenders, leading them to invest less in the sequential-move games compared to its simultaneous counterpart, particularly when free-riding among defenders are moderate. However, severe free-riders may have a diminished ability to do so in a sequential game and are forced to invest more. We also showed that the relative decrease of security investments in the sequential game compared to that of simultaneous game aligns with the notion of alpha-centrality in the reduced network.

8 Acknowledgement

This work is supported by the NSF under grants CNS-2012001 and CCF-2144283, and by the ARO under contract W911NF1810208.

References

[1] *5 Ways Cybercriminals are as Professional & Advanced as Your IT Security Teams*. en-US. Section: hybrid IT. Jan. 2022. URL: https://zivaro.com/cybercriminals-as-advanced-as-your-it-security-teams/ (visited on 06/06/2023).

[2] M. S. Bartlett. "An Inverse Matrix Adjustment Arising in Discriminant Analysis". In: *The Annals of Mathematical Statistics* 22.1 (Mar. 1951). Publisher: Institute of Mathematical Statistics, pp. 107–111. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177729698. URL: https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-22/issue-1/An-Inverse-Matrix-Adjustment-Arising-in-Discriminant-Analysis/10.1214/aoms/1177729698.full (visited on 03/05/2023).

[3] Phillip Bonacich and Paulette Lloyd. "Eigenvector-like measures of centrality for asymmetric relations". en. In: *Social Networks* 23.3 (July 2001), pp. 191–201. ISSN: 0378-8733. DOI: 10.1016/S0378-8733(01)00038-7. URL: https://www.sciencedirect.com/science/article/pii/S0378873301000387 (visited on 02/07/2023).

[4] Jerome Bracken and James T. McGill. "Defense Applications of Mathematical Programs with Optimization Problems in the Constraints". In: *Operations Research* 22.5 (1974). Publisher: INFORMS, pp. 1086–1096. ISSN: 0030-364X. URL: https://www.jstor.org/stable/169661 (visited on 01/05/2023).

[5] Jerome Bracken and James T. McGill. "Mathematical Programs with Optimization Problems in the Constraints". In: *Operations Research* 21.1 (1973). Publisher: INFORMS, pp. 37–44. ISSN: 0030-364X. URL: https://www.jstor.org/stable/169087 (visited on 01/05/2023).

[6] Yann Bramoullé and Rachel Kranton. "Public goods in networks". en. In: *Journal of Economic Theory* 135.1 (July 2007), pp. 478–494. ISSN: 0022-0531. DOI: 10.1016/j.jet.2006.06.006. URL: https://www.sciencedirect.com/science/article/pii/S0022053106001220 (visited on 02/22/2023).

[7] Yann Bramoullé, Rachel Kranton, and Martin D'Amours. "Strategic Interaction and Networks". en. In: *American Economic Review* 104.3 (Mar. 2014), pp. 898–930. ISSN: 0002-8282. DOI: 10.1257/aer.104.3.898. URL: https://www.aeaweb.org/articles?id=10.1257/aer.104.3.898 (visited on 12/13/2022).

[8] J P Morgan Chase. "How Criminals Use Social Engineering to Target Your Company". en. In: 10 ().

[9] Matthew Elliott and Benjamin Golub. "A Network Approach to Public Goods". In: *Journal of Political Economy* 127.2 (Apr. 2019). Publisher: The University of Chicago Press, pp. 730–776. ISSN: 0022-3808. DOI: 10.1086/701032. URL: https://www.journals.uchicago.edu/doi/abs/10.1086/701032 (visited on 02/22/2023).

[10] Ido Erev and Amnon Rapoport. "Provision of Step-Level Public Goods: The Sequential Contribution Mechanism". In: *The Journal of Conflict Resolution* 34.3 (1990). Publisher: Sage Publications, Inc., pp. 401–425. ISSN: 0022-0027. URL: https://www.jstor.org/stable/174222 (visited on 02/19/2023).

[11] Jens Grossklags, Nicolas Christin, and John Chuang. "Secure or insure?: a game-theoretic analysis of information security games". en. In: *Proceedings of the 17th international conference on World Wide Web*. Beijing China: ACM, Apr. 2008, pp. 209–218. ISBN: 978-1-60558-085-2. DOI: 10.1145/1367497.1367526. URL: https://dl.acm.org/doi/10.1145/1367497.1367526 (visited on 02/22/2023).

[12] *How cyber attacks work*. en. URL: https://www.ncsc.gov.uk/information/how-cyber-attacks-work (visited on 06/06/2023).

[13] Libin Jiang, Venkat Anantharam, and Jean Walrand. "How Bad Are Selfish Investments in Network Security?" In: *IEEE/ACM Transactions on Networking* 19.2 (Apr. 2011). Conference Name: IEEE/ACM Transactions on Networking, pp. 549–560. ISSN: 1558-2566. DOI: 10.1109/TNET.2010.2071397.

[14] Kun Jin and Mingyan Liu. "Multi-planner Intervention in Network Games with Community Structures". In: *2021 60th IEEE Conference on Decision and Control (CDC)*. ISSN: 2576-2370. Dec. 2021, pp. 2374–2379. DOI: 10.1109/CDC45484.2021.9683518.

[15] Kun Jin et al. "Network Games with Strategic Machine Learning". en. In: *Decision and Game Theory for Security*. Ed. by Branislav Bošanský et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 118–137. ISBN: 978-3-030-90370-1. DOI: 10.1007/978-3-030-90370-1_7.

[16] Dmytro Korzhyk et al. "Stackelberg vs. Nash in Security Games: An Extended Investigation of Interchangeability, Equivalence, and Uniqueness". In: *Journal of Artificial Intelligence Research* 41 (June 2011). arXiv:1401.3888 [cs], pp. 297–327. ISSN: 1076-9757. DOI: 10.1613/jair.3269. URL: http://arxiv.org/abs/1401.3888 (visited on 03/02/2023).

[17] Nir Kshetri. “The simple economics of cybercrimes”. In: *Security & Privacy, IEEE* 4 (Feb. 2006), pp. 33–39. DOI: 10.1109/MSP.2006.27.

[18] Ankur A. Kulkarni and Uday V. Shanbhag. “An Existence Result for Hierarchical Stackelberg v/s Stackelberg Games”. In: *IEEE Transactions on Automatic Control* 60.12 (Dec. 2015). Conference Name: IEEE Transactions on Automatic Control, pp. 3379–3384. ISSN: 1558-2523. DOI: 10.1109/TAC.2015.2423891.

[19] Howard Kunreuther and Geoffrey Heal. “Interdependent Security”. In: *Journal of Risk and Uncertainty* 26.2/3 (2003). Publisher: Springer, pp. 231–249. ISSN: 0895-5646. URL: https://www.jstor.org/stable/41755017 (visited on 02/15/2023).

[20] Yves Zenou Matthew O. Jackson. “Games on Networks”. en. In: ().

[21] R. Ann Miura-Ko et al. “Security Decision-Making among Interdependent Organizations”. In: *2008 21st IEEE Computer Security Foundations Symposium*. ISSN: 2377-5459. June 2008, pp. 66–80. DOI: 10.1109/CSF.2008.25.

[22] Parinaz Naghizadeh and Mingyan Liu. *Provision of Public Goods on Networks: On Existence, Uniqueness, and Centralities*. arXiv:1604.08910 [cs]. May 2016. DOI: 10.48550/arXiv.1604.08910. URL: http://arxiv.org/abs/1604.08910 (visited on 11/21/2022).

[23] Hans-Theo Normann. “Simultaneous and Sequential Contributions to Step-Level Public Goods: One vs. Two Provision Levels”. en. In: ().

[24] Jong-shi Pang and Masao Fukushima. “Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games”. English. In: *Computational Management Science* 2.1 (Jan. 2005). Num Pages: 21-56 Publisher: Springer Nature B.V., pp. 21–56. ISSN: 1619697X. DOI: 10.1007/s10287-004-0010-0. URL: https://www.proquest.com/docview/612782537/abstract/5A753DC80F4945B7PQ/1 (visited on 12/30/2022).

[25] Francesca Parise and Asuman Ozdaglar. *A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis*. arXiv:1712.08277 [physics]. Aug. 2018. DOI: 10.48550/arXiv.1712.08277. URL: http://arxiv.org/abs/1712.08277 (visited on 01/04/2023).

[26] Gesualdo Scutari et al. “Real and Complex Monotone Communication Games”. In: *IEEE Transactions on Information Theory* 60.7 (July 2014). Conference Name: IEEE Transactions on Information Theory, pp. 4197–4231. ISSN: 1557-9654. DOI: 10.1109/TIT.2014.2317791.

[27] Hanif D. Sherali. “A Multiple Leader Stackelberg Model and Analysis”. In: *Operations Research* 32.2 (1984). Publisher: INFORMS, pp. 390–404. ISSN: 0030-364X. URL: https://www.jstor.org/stable/172107 (visited on 01/04/2023).

[28] Michael Tsatsomeros. “Lecture Notes on Matrices with Positive Principal Minors: Theory and Applications”. en. In: (), p. 31.

[29] Hal Varian. "System Reliability and Free Riding". en. In: *Economics of Information Security*. Ed. by L. Jean Camp and Stephen Lewis. Advances in Information Security. Boston, MA: Springer US, 2004, pp. 1–15. ISBN: 978-1-4020-8090-6. DOI: 10.1007/1-4020-8090-5_1. URL: https://doi.org/10.1007/1-4020-8090-5_1 (visited on 02/22/2023).

[30] Hal R. Varian. "Sequential contributions to public goods". en. In: *Journal of Public Economics* 53.2 (Feb. 1994), pp. 165–186. ISSN: 0047-2727. DOI: 10.1016/0047-2727(94)90019-1. URL: https://www.sciencedirect.com/science/article/pii/0047272794900191 (visited on 02/15/2023).

[31] Wei Wei. "Tutorials on Advanced Optimization Methods". In: *arXiv: Optimization and Control* (July 2020). URL: https://www.semanticscholar.org/paper/Tutorials-on-Advanced-Optimization-Methods-Wei/ca78817bad62535e8aec8e070b11ee20ec59cbe9 (visited on 12/31/2022).

[32] Fuzhen Zhang, ed. *The Schur Complement and Its Applications*. en. Vol. 4. Numerical Methods and Algorithms. New York: Springer-Verlag, 2005. ISBN: 978-0-387-24271-2. DOI: 10.1007/b105056. URL: http://link.springer.com/10.1007/b105056 (visited on 03/02/2023).