

Preventive and Reactive Cyber Defense Dynamics with Ergodic Time-dependent Parameters Is Globally Attractive

Yujuan Han, Wenlian Lu, *Senior Member, IEEE*, Shouhuai Xu

Abstract—Cybersecurity dynamics is a mathematical approach to modeling and analyzing cyber attack-defense interactions in networks. In this paper, we advance the state-of-the-art in characterizing one kind of cybersecurity dynamics, known as *preventive and reactive cyber defense dynamics*, which is a family of highly nonlinear system models. We prove that this dynamics in its general form with *time-dependent* parameters is *globally attractive* when the time-dependent parameters are *ergodic*, and is *(almost) periodic* when the time-dependent parameters have the stronger properties of being *(almost) periodic*. Our results supersede the state-of-the-art ones, including that the same type of dynamics but with *time-independent* parameters is *globally convergent*.

Index Terms—Cybersecurity dynamics, preventive and reactive cyber defense dynamics, global attractivity, network science

I. INTRODUCTION

Cyberspace is a complex system that has become a critical infrastructure. However, our understanding of its security is still superficial, explaining why there are so many cyber attacks on a daily basis. This calls for research in understanding cybersecurity at many levels of abstractions, ranging from macroscopic to microscopic [1, 2, 3]. At a macroscopic level, Kephart and White [4, 5] adapt the classic biological epidemic models [6] to cyberspace, while inheriting the homogeneity assumption that each node in a network has equal chance in attacking any other node. This approach is later extended to accommodate heterogeneous network structures represented by *arbitrary* adjacency matrices [7].

These studies [4, 5, 7] have inspired a systematic approach, dubbed Cybersecurity Dynamics [1, 2, 3], which can be characterized as follows. First, it proposes using arbitrary matrices to describe the *attack-defense* structures that are induced by cybersecurity (e.g., access control) policies enforced on top of physical networks. That is, adjacency matrices are used to model which nodes are allowed to communicate with which other nodes through some routing paths (with each typically consisting of multiple point-to-point physical communication links), rather than modeling the physical links.

Second, computer epidemic models, including [4, 5, 7, 8] and their numerous follow-up studies, often focus on investigating *epidemic threshold*, which distinguishes the parameter

regime in which the spreading dies out from the parameter regime in which the spreading doesn't. While important, this understanding is far from sufficient in cybersecurity. For example, we need to know whether the spreading will be converging or not when it does not die out.

Third, the rich semantics of cyber attack-defense interactions has led to families of cybersecurity dynamics models, including: preventive and reactive cyber defense dynamics [9, 10, 11, 12, 13], active cyber defense dynamics [14, 15, 16], adaptive cyber defense dynamics [17, 18], and proactive cyber defense dynamics [19]. These theoretical studies have deepened our understanding of cybersecurity. For example, now we know: preventive and reactive cyber defense dynamics is *globally convergent* in certain settings [12, 13], and *global convergence* is a nice cybersecurity property that makes it possible to predict and manage cybersecurity [11]; in contrast, active cyber defense dynamics can be Chaotic [16].

In this paper, we focus on the aforementioned preventive and reactive cyber defense dynamics, which is a family of highly nonlinear system models that are initiated in [9] and inspired by [7]. This dynamics aims to model the interactions between two classes of cyber attacks and two classes of cyber defenses. The two classes of attacks are: *push-based* attacks (including computer malware spreading) and *pull-based* attacks (including “drive-by download” attacks; i.e., a computer gets compromised when visiting a malicious website [20]). The two classes of defenses are: *preventive* defenses, which include the use of access control and intrusion prevention mechanisms to attempt to prevent attacks from succeeding; and *reactive* defenses, which include the use of anti-malware and intrusion detection mechanisms to attempt to detect and clean up compromised computers. Since this dynamics uses a certain *product* term, which will be elaborated later, to model the collective effect on a node when attacked by others, it is also known as the \prod -model.

Closely related to the \prod -model is the the N -intertwined model [21], or the \sum -model because it uses a certain *additive* term (which will be elaborated later as well) to model the collective effect on a node when attacked by others. This model is also inspired by [7] and has been studied in, for example, [22, 23, 24, 25, 26].

A more general preventive and reactive cyber defense dynamics is introduced in [13], which accommodates the aforementioned \prod -model and \sum -model as two special cases and is thus dubbed the *unified dynamics* or *unified model*. A remarkable result is that the unified model is globally convergent in

Y. Han is with College of Information Engineering, Shanghai Maritime University, China and Department of Computer Science, University of Texas at San Antonio, USA. W. Lu is with School of Mathematical Sciences, Fudan University, China; Shanghai Center for Mathematical Sciences, Fudan University, China; and Shanghai Key Laboratory for Contemporary Applied Mathematics, China. S. Xu is with Department of Computer Science, University of Texas at San Antonio, USA. Correspondence: shxu@cs.utsa.edu.

the *entire* parameter universe [13]. As shown in [13], this result supersedes many results presented in the literature, which often deal with some special cases of the unified model. While fairly general, this unified model [13] only accommodates *time-independent* parameters (i.e., parameter values do not change over time). This *time-independence* of parameters is restrictive, and should be eliminated to accommodate *time-dependent* parameters to make the theoretical results more widely applicable. This motivates the present study.

A. Our Contributions

We investigate preventive and reactive cyber defense dynamics with *time-dependent* parameters, or the *unified model with time-dependent parameters*. Our results supersede the ones obtained in the unified model with time-independent parameters [13]; our results also supersede previous results that are obtained in special cases of the time-dependent \prod -model [18, 27, 28, 29] and the time-dependent \sum -model [30, 31]. All of these results typically deal with convergence to an equilibrium, meaning that their techniques are no longer applicable in our setting, explaining why we adopt the skew-product semiflow approach and the multiplicative ergodic theorem in the present paper. We prove that the unified model with *ergodic* parameters is globally attractive (i.e., converging to a unique trajectory regardless of the initial value as illustrated in Figure 1(a) where the y -axis represents a metric of interest (e.g., the fraction of compromised nodes in a network [32]), whereas Figure 1(b) illustrates the absence of global attractivity. We further prove that when the parameters are almost periodic (vs. periodic), the unified model with time-dependent parameters is also almost periodic (correspondingly, periodic). In addition, we present bounds on the globally attractive trajectory, which are useful because (for example) the upper bound can be seen as the worst case scenario for decision-making purposes when only partial information about the parameters is given.

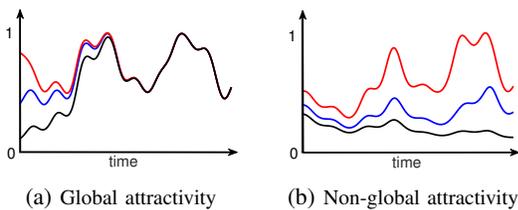


Fig. 1: Illustration of global attractivity vs. non-global attractivity, where the latter is sensitive to the initial value and may be even Chaotic and therefore unpredictable.

In order to systematize knowledge, we use Figure 2 to highlight how the present paper supersedes the literature results. Our Theorem 4 supersedes the global convergence of the unified model with *time-independent* parameters presented in [13], which in turn supersedes numerous prior results reported in the literature; this is because the ergodicity condition (or assumption) required by Theorem 4 naturally holds in the time-independent setting. When there are no pull-based attacks (i.e., considering push-based attacks only), our Theorem 3 shows that under even weaker condition (than ergodicity)

the unified model with *time-dependent* parameters globally converge to the special equilibrium $\mathbf{0}$ (i.e., there are no compromised nodes); this supersedes the literature results on the \prod -dynamics and \sum -dynamics with *time-dependent* parameters, because the condition required by our Theorem 3 naturally holds in the \prod -model and \sum -model with time-dependent parameters investigated in the literature.

B. Related Work

To the best of our knowledge, there is no prior study that aims to characterize the unified model with time-dependent parameters in the entire parameter universe. Prior studies related to preventive and reactive cyber defense dynamics can be divided into two categories: considering *time-independent* parameters vs. considering *time-dependent* parameters. For *time-independent* models, the state-of-the-art is the global convergence result of the unified model [13], which supersedes numerous results in the \prod -model (e.g., [11, 12, 13]) and the \sum -model (e.g., [21, 22, 23, 24, 25, 26]). Time-dependent parameters have been investigated in the \prod -model and \sum -model, but only to the extent of understanding when the dynamics converges to the special equilibrium $\mathbf{0}$ while assuming there are no pull-based attacks. Specifically, [18] investigates the \prod -model with time-dependent parameters and identifies a sufficient condition under which the dynamics converges to $\mathbf{0}$; [27, 28, 29] investigate the convergence to $\mathbf{0}$ of the \prod -model with time-independent parameters. For the \sum -model with time-dependent parameters while assuming there are no pull-based attacks, sufficient conditions under which the dynamics converges to $\mathbf{0}$ are presented in [30, 31, 33, 34].

C. Paper Outline

Section II reviews some mathematical knowledge. Section III presents and investigates the unified model with time-dependent parameters. Section IV reports our simulation study. Section V concludes the paper with open problems.

II. MATHEMATICAL PRELIMINARIES

Let \mathbb{R} be the set of real numbers, $\mathbb{R}_{\geq 0}^n = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_v \geq 0 \text{ for } 1 \leq v \leq n\}$, $\mathbb{R}_{> 0}^n = \{x \in \mathbb{R}_{\geq 0}^n : x_v > 0 \text{ for some } 1 \leq v \leq n\}$, $\mathbb{R}_{\gg 0}^n = \{x \in \mathbb{R}_{> 0}^n : \forall v, x_v > 0\}$, $[0, 1]^n = \{x \in \mathbb{R}^n : 0 \leq x_v \leq 1\}$, and $[0, 1]^{n,n} = \{A \in \mathbb{R}^{n \times n} : 1 \leq A_{uv} \leq 0\}$. Let I_n denote the $n \times n$ identity matrix. For a matrix $A = (a_{vu})$, let $A \geq \delta$ denote that every element of A is greater than δ , namely $a_{vu} \geq \delta \forall v, u$. For two vectors $x, z \in \mathbb{R}^n$, let $x \geq z$ denote $x - z \in \mathbb{R}_{\geq 0}^n$, $x > z$ denote $x - z \in \mathbb{R}_{> 0}^n$, and $x \gg z$ denote $x - z \in \mathbb{R}_{\gg 0}^n$. Table I summarizes the other notations.

A. Ergodicity and Almost Periodicity

We propose modeling *time-dependent* parameters $\{y(t)\}_{t \geq 0}$ as a sequence drawn from an *ergodic* stochastic process because we can only observe a single sequence of a stochastic process in the real world. Therefore, it is necessary to require that the observed sequence is representative of the stochastic

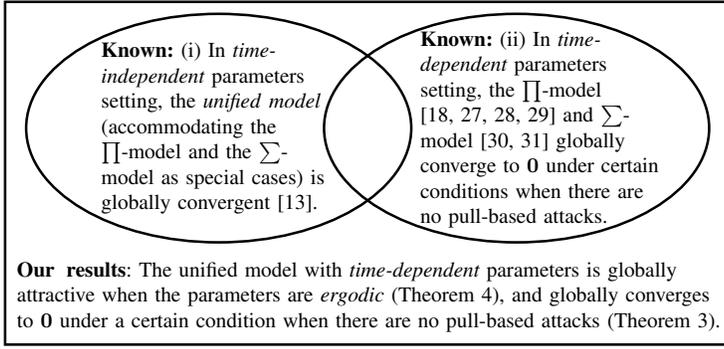


TABLE I: Notations used in the paper.

$G(t), A(t)$	attack-defense structure $G(t) = (V, E(t))$ where $V = \{1, \dots, n\}$ is the set of nodes and $E(t)$ is the set of arcs; its adjacency matrix $A(t) = [a_{vu}(t)]_{n \times n}$ where $a_{vu}(t) = 1$ if and only if $(u, v) \in E(t)$
$N_v(t)$	v 's neighbors that are allowed to communicate with v at time t , i.e., $N_v(t) = \{u \in V : (u, v) \in E(t)\}$
$\gamma_{vu}(t), \Gamma(t)$	$\gamma_{vu}(t) \in [0, 1]$ is the probability a <i>secure</i> node $v \in V$ becomes <i>compromised</i> because of the push-based attack waged by <i>compromised</i> neighbor $u \in N_v(t)$; $\Gamma(t) = [\gamma_{vu}(t)]_{n \times n}$
$\mu(C(t))$	the maximum Lyapunov exponent (MLE) of system $dz(t)/dt = C(t)z(t)$; $\mu(C(t)) = \limsup_{t \rightarrow \infty} \frac{1}{t} \log \ U(t, 0)\ $, where $U(t, 0)$ is the <i>fundamental solution matrix</i> of the system
$\lambda_1(B), \rho(B)$	the maximum eigenvalue of B (in real part); the spectral radius of B
$[x]_{\mathcal{V}}, [B]_{\mathcal{V}, \mathcal{W}}$	for any vector $x \in \mathbb{R}^n$ and any index subset $\mathcal{V} \subseteq \{1, \dots, n\}$, $[x]_{\mathcal{V}} = (x_v)_{v \in \mathcal{V}} \in \mathbb{R}^{ \mathcal{V} }$; for any matrix $B = [b_{vu}] \in \mathbb{R}^{n \times n}$ and two index subsets $\mathcal{V}, \mathcal{W} \subseteq \{1, \dots, n\}$, $[B]_{\mathcal{V}, \mathcal{W}} = [b_{vu}]_{v \in \mathcal{V}, u \in \mathcal{W}} \in \mathbb{R}^{ \mathcal{V} \times \mathcal{W} }$
$\mathbf{0}, \mathbf{0}_d$	$\mathbf{0} = [0, \dots, 0] \in \mathbb{R}^n$, $\mathbf{0}_d = [0, \dots, 0] \in \mathbb{R}^d$
$i_v(t), s_v(t)$	the probability that node $v \in V$ is <i>compromised</i> and <i>secure</i> at time t , where $i_v(t) + s_v(t) = 1$
$\alpha_v(t), \alpha(t)$	$\alpha_v(t) \in [0, 1]$ is the probability that a <i>secure</i> node $v \in V$ becomes <i>compromised</i> at time t because of pull-based attacks; $\alpha(t) = [\alpha_1(t), \dots, \alpha_n(t)]$
$\beta_v(t), \beta(t)$	$\beta_v(t) \in [0, 1]$ is the probability that a <i>compromised</i> node $v \in V$ becomes <i>secure</i> at time t because of the reactive defense; $\beta(t) = [\beta_1(t), \dots, \beta_n(t)]$
$y(t)$	$y(t) = [\alpha(t), \beta(t), \text{vec}(A(t))^T, \text{vec}(\Gamma(t))^T]$ is the vector of model parameters (including the attack-defense structure), where $\text{vec}(A(t))$ is the column vector obtained by stacking the columns of matrix $A(t)$
$D_x f(x, y)$	$D_x f(x, y) = [\partial f_v(x, y) / \partial x_u]_{u, v \in V}$, the Jacobian matrix of $f : X \times Y \rightarrow X$ at point (x, y) w.r.t. x
$\mathcal{M}(\mathbb{R}, \mathbb{R}^m)$	the space of measurable functions from \mathbb{R} to \mathbb{R}^m

process, meaning that the averaged behavior of the sequence is the same as the average over the probability space. Formally,

Definition 1 (ergodicity and mean value [35, 36]). *Let $Y \subseteq \mathcal{M}(\mathbb{R}, \mathbb{R}^m)$ be a compact space of measurable functions from \mathbb{R} to \mathbb{R}^m , $(Y, \mathcal{F}, \mathbb{P})$ be a probability space, and $\theta : \mathbb{R}^+ \times Y \rightarrow Y$ be the right-shift translation that $\theta(s, y(t)) = y(t + s)$, $\forall s, t \geq 0$. Measure \mathbb{P} is called θ -invariant if for any $K \in \mathcal{F}$, $\mathbb{P}(K) = \mathbb{P}(\theta^{-1}(t)K)$ holds for all $t \in \mathbb{R}$, and called θ -ergodic if for any $A \in \mathcal{F}$ that satisfies $A = \theta^{-1}(t)A$, $\forall t \in \mathbb{R}$, A has measure 0 or 1. We call $y \in Y$, namely $\{y(t)\}_{t \geq 0} \in Y$, ergodic if and only if \mathbb{P} is θ -invariant and θ -ergodic.*

For any ergodic process $\{y(t)\}_{t \geq 0}$, $M(y)$ is called the mean

value of $\{y(t)\}_{t \geq 0}$, where

$$M(y) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_a^{a+T} y(\tau) d\tau \quad (1)$$

holds uniformly with respect to any a almost surely.

Many stochastic processes are ergodic, such as a sequence of independent and identically distributed random variables and ergodic Markov processes [37]. Moreover, (almost) periodic functions are also ergodic [35]. Formally, we have:

Definition 2 (almost periodicity [38]). *A continuous function $y(t) : \mathbb{R} \rightarrow \mathbb{R}^m$ is said almost periodic in t if for any $\epsilon > 0$, there exists a number $l(\epsilon) > 0$ such that every interval of length $l(\epsilon)$ contains a point $\xi \in \mathbb{R}$ and*

$$\|y(t + \xi) - y(t)\| < \epsilon, \quad \forall t \in \mathbb{R},$$

where ξ is called “an ϵ -translation number of $y(t)$ ”. A continuous function $f(t, x) : \mathbb{R} \times X \rightarrow \mathbb{R}^n$ is called almost periodic in t if for any $x \in X$, $f(t, x)$ is almost periodic in t .

An almost periodic time-dependent parameter $g(t)$, if not periodic, has no period, meaning $g(t + \tau) = g(t), \forall t \in \mathbb{R}$ does not hold for any $\tau \in \mathbb{R}$, but $g(t + \tau) \approx g(t), \forall t \in \mathbb{R}$ holds with any degree of approximation for infinitely many $\tau \in \mathbb{R}$, where τ may be very large. For example, $g(t) = 0.2 \times \sin(\pi t) + 0.1 \times \sin(2\sqrt{2}\pi t) + 0.3$, which contains three periodic terms, is almost periodic but not periodic.

B. Subhomogeneous and Cooperative Dynamical Systems

We will take advantage of cooperative and subhomogeneous dynamical systems, where the former means that all of the off-diagonal terms of the Jacobian matrix of a dynamical system are non-negative and the latter (or sublinearity) is a generalization of concavity.

Definition 3 (cooperative dynamical system [39]). *Consider a region $X \subseteq \mathbb{R}_{\geq 0}^n$, a subspace $Y \subseteq \mathcal{M}(\mathbb{R}, \mathbb{R}_{\geq 0}^m)$, $x = [x_1, \dots, x_n]^T \in X$, $y \in Y$, and $f(x, y) = [f_1(x, y), \dots, f_n(x, y)]^T : X \times Y \rightarrow X$. A nonautonomous system*

$$\frac{dx}{dt} = f(x, y(t)) \quad (2)$$

is said to be cooperative if $\partial f_v(x, y) / \partial x_u \geq 0$ holds for $\forall u \neq v$ and $\forall (x, y) \in X \times Y$.

Definition 4 (subhomogeneity and monotonicity [40]). A continuous map $f(x, y) : X \times Y \rightarrow X$ is said to be

- subhomogeneous if $f(\eta x, y) \geq \eta f(x, y)$ holds for any $x \in X$, $y \in Y$, and $\eta \in (0, 1)$.
- strictly subhomogeneous if $f(\eta x, y) > \eta f(x, y)$ holds for any $x \in X$ with $x \gg 0$, any $y \in Y$ and $\eta \in (0, 1)$.
- strongly subhomogeneous if $f(\eta x, y) \gg \eta f(x, y)$ holds for any $x \in X$ with $x \gg 0$, any $y \in Y$ and $\eta \in (0, 1)$.
- monotone if $f(x_1, y) \geq f(x_0, y)$ holds for any $x_1 \geq x_0$ and $y \in Y$.
- strictly monotone if $f(x_1, y) > f(x_0, y)$ holds for any $x_1 > x_0$ and $y \in Y$.
- strongly monotone if $f(x_1, y) \gg f(x_0, y)$ holds for any $x_1 > x_0$ and $y \in Y$.

C. Globally Attractive Dynamical Systems

The concept of *uniform persistence* describes the behavior that trajectories are eventually uniformly away from the boundary of some closed invariant subset [41, 42]. Intuitively, when the origin $\mathbf{0}$ is on the boundary, uniform persistence implies the instability of the origin $\mathbf{0}$.

Definition 5 (persistence and uniform persistence [42]). Consider a closed region $X \subset \mathbb{R}_{\geq 0}^n$ with $\mathbf{0} \in X$ and a function space $Y \subset \mathcal{M}(\mathbb{R}, \mathbb{R}^m)$. For given $x_0 \in X \setminus \{\mathbf{0}\}$ and $y_0 \in Y$, a continuous map $\psi(t, x_0, y_0) : \mathbb{R} \rightarrow X$ is said to be

- persistent if $\liminf_{t \rightarrow +\infty} \psi(t, x_0, y_0) \gg \mathbf{0}$ holds.
- uniformly persistent if $\liminf_{t \rightarrow +\infty} \psi(t, x_0, y) \gg \mathbf{0}$ holds for $\forall x_0 \in X \setminus \{\mathbf{0}\}$ and $\forall y \in Y$.

Denote by $\psi(t, x_0, y)$ the solution to system (2) with respect to initial value $x(0) = x_0$ and $y \in Y$ (representing time-dependent parameters in the context of the present paper). Now we introduce the the definition of *global attractivity*.

Definition 6 (Global attractivity). Consider system (2) with compact spaces $X \subseteq \mathbb{R}_{\geq 0}^n$ and $Y \subseteq \mathcal{M}(\mathbb{R}, \mathbb{R}^n)$, where $(Y, \mathcal{F}, \mathbb{P})$ is a probability space and $y = \{y(t)\}_{t \geq 0} \in Y$ is a sample (or realization) of some time-dependent parameters.

- For system (2), equilibrium $\mathbf{0}$ is said to be almost surely globally attractive if $\lim_{t \rightarrow \infty} \psi(t, x_0, y) = \mathbf{0}$ holds for any $x_0 \in X$ and almost every $y \in Y$.
- For system (2) with a given $y \in Y$, a trajectory $\psi(t, x_y^*, y)$ is said to be globally attractive if $\lim_{t \rightarrow \infty} \|\psi(t, x_0, y) - \psi(t, x_y^*, y)\| = 0$ holds for any $x_0 \in X \setminus \{\mathbf{0}\}$.
- System (2) is said to be almost surely globally attractive if for almost every $y \in Y$, there exists a globally attractive trajectory $\psi(t, x_y^*, y)$, namely that $\lim_{t \rightarrow \infty} \|\psi(t, x_0, y) - \psi(t, x_y^*, y)\| = 0$ holds for any $x_0 \in X \setminus \{\mathbf{0}\}$.

Theorem 1 below bridges the preceding two concepts.

Theorem 1 (Theorem 2.3.5 in [40]). Let $X \subseteq \mathbb{R}_{\geq 0}^n$ and Y be compact spaces, where Y has no nonempty, proper, closed invariant subset with respect to θ and there is a metric d such that for any distinct $y_1, y_2 \in Y$, $\inf_{t \in \mathbb{R}} d(\theta(t, y_1), \theta(t, y_2)) >$

0. Let $H : \mathbb{R} \times X \times Y \rightarrow X \times Y$ be the skew-product semi-flow associated to system (2) in the form

$$H(t, x, y) = (\psi(t, x, y), \theta(t, y)),$$

where $\theta : \mathbb{R} \times Y \rightarrow Y$ is the right-shift translation (see Definition 1). Denote by $\text{int}(X) = X \cap \mathbb{R}_{\gg 0}^n$. Suppose

- for any $\eta \in (0, 1)$, $x_1, x_2 \in \text{int}(X)$, $\eta x_1 \leq x_2 \leq \eta^{-1} x_1$ implies that $\eta \psi(t, x_1, y) \leq \psi(t, x_2, y) \leq \eta^{-1} \psi(t, x_1, y)$, $\forall t \in \mathbb{R}, \forall y \in Y$; and
- there exists $t_0 > 0$ and $y_0 \in Y$ such that for any $\eta \in (0, 1)$, $x_1, x_2 \in \text{int}(X)$, $\eta x_1 \leq x_2 \leq \eta^{-1} x_1$ implies that $\eta \psi(t_0, x_1, y_0) \ll \psi(t_0, x_2, y_0) \ll \eta^{-1} \psi(t_0, x_1, y_0)$.

If H has a compact ω -limit set $K_0 \subset \text{int}(X) \times Y$, then the natural projection $p : X \times Y \rightarrow Y$ is a flow isomorphism restricted on $H(t, \cdot, \cdot) : K_0 \times Y \rightarrow K_0$ and for every compact ω -limit set $w(x, y) \subset \text{int}(X) \times Y$, we have $w(x, y) = K_0$ and $\lim_{t \rightarrow \infty} \|\psi(t, x, y) - \psi(t, x^*, y)\| = 0$, where $(x^*, y) = K_0 \cap p^{-1}(y)$.

D. Time-Dependent Linear Cooperative Systems

We will leverage time-dependent linear systems and δ -matrices [43]. Consider a time-dependent linear system

$$\frac{dz}{dt} = B(t)z(t), \quad (3)$$

where $z(t) \in \mathbb{R}^n$ and $B(t) = [b_{ij}(t)]_{i,j=1}^n$ with $b_{ij}(t) \geq 0$ for $i \neq j$ and $t \in \mathbb{R}$. We will relate $B(t)$ to a time-dependent (attack-defense) graph $\mathcal{G}(B(t)) = (V, E(t), B(t))$ with node set $V = \{1, \dots, n\}$, time-dependent arc set $E(t)$ and time-dependent weight matrix $B(t)$. A matrix $B^\delta(t) = [b_{ij}^\delta(t)]$ with

$$b_{ij}^\delta(t) = \begin{cases} b_{ij}(t) & b_{ij}(t) \geq \delta \text{ \& } i \neq j \\ 0 & b_{ij}(t) < \delta \text{ \& } i \neq j \\ b_{ii}(t) & i = j \end{cases}$$

is called a δ -matrix of $B(t)$ and its associated graph $\mathcal{G}(B^\delta(t))$ is called δ -graph of $B(t)$.

When $\{B(t)\}_{t \geq 0}$ is ergodic, meaning system (3) is ergodic, let $M(B)$ denote its mean value (cf. Definition 1). We will use the following lemma, which is a basic result regarding ergodic time-dependent linear system (3). This lemma says that the fundamental solution matrix of system (3) is positive when $\{B(t)\}_{t \in \mathbb{R}}$ is ergodic and the associated graph $\mathcal{G}(M(B))$ is strongly connected; its proof is deferred to Appendix A.

Lemma 1. Consider system (3) with bounded and ergodic $\{B(t)\}_{t \geq 0}$. Let $M(B)$ denote the mean value of $\{B(t)\}_{t \geq 0}$ and $U(t, s)$ the fundamental solution matrix of system (3). Then, we have:

- every element of $U(t, s)$ is nonnegative for any $t \geq s \geq 0$; and
- if $\mathcal{G}(M(B))$ is strongly connected, there exists $T > 0$ such that for each $\Delta \geq T$, one can find $\epsilon > 0$ (dependent on Δ) such that every element of $U(s + \Delta, s)$ is greater than ϵ for any $s \geq 0$.

III. MODEL AND ANALYSIS

The preventive and reactive cyber defense dynamics model introduced in [13], dubbed unified model with time-independent parameters, unifies two families of models into a single framework. The unified dynamics or model is proven to be *globally convergent* (i.e., always converging to some equilibrium) [13]. Now we present and analyze its extension to *unified model with time-dependent parameters*.

A. The Unified Model with Time-Dependent Parameters

The unified model with time-dependent parameters still describes the dynamics of the global cybersecurity state incurred by the interactions between two classes of cyber attacks (i.e., pull-based attacks and push-based attacks) and two classes of defenses (i.e., preventive defenses and reactive defenses) in a network. Intuitively, the attack-defense structure incurred by the attack-defense interactions at time t can be described by a directed graph, denoted by $\tilde{G}(t) = (\tilde{V}(t), \tilde{E}(t))$, where $\tilde{V}(t)$ is the node (representing a computer) set and $\tilde{E}(t)$ is the arc set, and $(u, v) \in \tilde{E}(t)$ means node u can communicate with, and therefore can wage push-based attacks against, node v at time t . As we will justify later, it suffices to consider a *time-independent* node set $V = \{1, \dots, n\}$, namely $G(t) = (V, E(t))$ rather than $\tilde{G}(t) = (\tilde{V}(t), \tilde{E}(t))$, because the evolution of $\tilde{V}(t)$ can be “encoded” into the evolution of $E(t)$, leading to a simpler representation $G(t) = (V, E(t))$ of attack-defense structures.

The effect of attacks against preventive defenses is modeled as follows. Push-based attacks take place on attack-defense structures $G(t) = (V, E(t))$. Denote by $A(t)$ the adjacency matrix of $G(t)$, where $a_{vu}(t) = 1$ if $(u, v) \in E(t)$ and $a_{vu}(t) = 0$ otherwise. Let $\gamma_{vu}(t) \in [0, 1]$ denote the probability that a push-based attack, which is waged by a compromised node $u \in V$ against a secure $v \in V$ over arc $(u, v) \in E(t)$ at time t , succeeds (i.e., causing v to become compromised); that is, $1 - \gamma_{vu}(t)$ represents the effectiveness of the preventive defense mechanism deployed at node v and/or the arc (u, v) . Let $\Gamma(t) = [\gamma_{vu}(t)]_{n \times n}$ denote the probability matrix corresponding to the adjacency matrix $A(t)$. The effect of push-based attacks over $G(t)$ can be described by $(G(t), \Gamma(t))$. On the other hand, pull-based attacks can be described by using $\alpha_v(t) \in [0, 1]$ to denote the probability that a secure node $v \in V$ becomes compromised at time t because of pull-based attacks; that is, $1 - \alpha_v(t)$ represents the effectiveness of preventive defense against pull-based attacks.

For modeling the effect of reactive defenses against successful attacks, let $\beta_v(t) \in [0, 1]$ denote the probability that a compromised node v is detected and “cleaned up” (i.e., becoming secure) at time t ; that is, $1 - \beta_v(t)$ represents the ineffectiveness of the reactive defense.

At any time t , a node $v \in V$ is in one of two cybersecurity states, *compromised* or *secure*. Let $i_v(t)$ and $s_v(t)$ respectively denote the probability that v is in the *compromised* state and the *secure* state at time t . Let $i(t) = [i_1(t), \dots, i_n(t)]$ and $s(t) = [s_1(t), \dots, s_n(t)]$. Figure 3 describes the state transition diagram of a node $v \in V$, where $h_v(i(t), \beta_v(t)) : [0, 1]^n \times [0, 1] \rightarrow [0, 1]$ is a function that outputs the probability

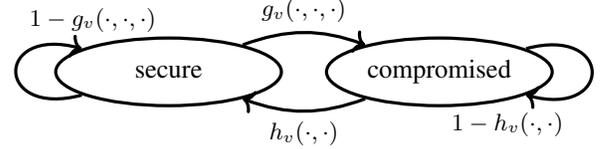


Fig. 3: State transition diagram of node $v \in V$ at time t .

that a compromised node v becomes secure at time t because of the deployed reactive defenses, and $g_v(i(t), \alpha_v(t), \Gamma(t)) : [0, 1]^n \times [0, 1] \times [0, 1]^{n \times n} \rightarrow [0, 1]$ is a function that outputs the probability that a secure node v becomes compromised at time t because of the push-based and pull-based attacks that penetrate the deployed preventive defenses. Figure 3 and the fact that $i_v(t) + s_v(t) = 1$ holds for any $v \in V$ and any $t \geq 0$, the unified model with time-dependent parameters is described by the following dynamical system for each $v \in V$:

$$\begin{aligned} \frac{di_v(t)}{dt} = & -h_v(i(t), \beta_v(t)) \cdot i_v(t) \\ & + g_v(i(t), \alpha_v(t), \Gamma(t)) \cdot (1 - i_v(t)). \end{aligned} \quad (4)$$

The research objective is to analyze system or model (4) with time-dependent parameters $\alpha_v(t)$, $\beta_v(t)$, $\gamma_{vu}(t)$ and $G(t) = (V, E(t))$. This is a challenging task because among other things, the term $g_v(i(t), \alpha_v(t), \Gamma(t))$ is highly nonlinear. In order to simplify presentation, we use $y(t)$ to denote the collection of parameters at time t , namely $y(t) = [\alpha(t), \beta(t), \text{vec}(A(t))^\top, \text{vec}(\Gamma(t))^\top]$, where $\text{vec}(\text{matrix})$ is the vector representation of a matrix *matrix* obtained by concatenating the columns; we use $f_v(i(t), y(t))$ to denote the right-hand side of system (4), namely

$$\begin{aligned} f_v(i(t), y(t)) = & -h_v(i(t), \beta_v(t)) \cdot i_v(t) \\ & + g_v(i(t), \alpha_v(t), \Gamma(t)) \cdot (1 - i_v(t)); \end{aligned} \quad (5)$$

and we let $f(i, y) = [f_1(i, y), \dots, f_n(i, y)]^\top$.

On the generality of model (4) in accommodating models studied in the literature. Model (4) is a general framework because h_v and g_v can be instantiated in many ways. In particular, model (4) degenerates to the unified model with time-independent parameters [13] by letting $G(t)$, $\gamma_{vu}(t)$, $\alpha_v(t)$ and $\beta_v(t)$ be time-independent. Moreover, the extended model also accommodates the aforementioned two models as special cases, namely the \prod -model because a \prod -term instantiates g_v as shown in Eq. (6) and the \sum -model because a \sum -term instantiates g_v as shown in Eq. (6).

$$\begin{aligned} & g_v(i(t), \alpha_v(t), \Gamma(t)) \quad (6) \\ = & \begin{cases} 1 - (1 - \alpha_v(t)) \prod_{u \in N_v(t)} (1 - \gamma_{vu}(t) i_u(t)), & \prod\text{-model} \\ \alpha_v(t) + \sum_{u \in N_v(t)} \gamma_{vu}(t) i_u(t), & \sum\text{-model} \end{cases} \end{aligned}$$

Both models set $h_v(i(t), \beta_v(t)) = \beta_v(t)$. Both models are nonlinear, especially the \prod -model, which is highly nonlinear and thus makes the unified model with time-dependent parameters highly nonlinear. These two models differ in how they

model the collective effect of push-based attacks waged by v 's neighbors against v , namely g_v .

How can $G(t) = (V, E(t))$ encode $\tilde{G}(t) = (\tilde{V}(t), \tilde{E}(t))$? Given $\tilde{G}(t) = (\tilde{V}(t), \tilde{E}(t))$ over time t , we can set $V = \bigcup_t \tilde{V}(t)$, meaning that a node $u \in V - \tilde{V}(t)$ can be treated as an isolated, dummy node in $G(t) = (V, E(t))$ such that $(u, \cdot) \notin E(t)$ and $(\cdot, u) \notin E(t)$ and $\alpha_u(t) = 0$. This justifies why it suffices to consider $G(t) = (V, E(t))$.

B. Some Properties of Functions h_v , g_v and Parameters $y(t)$

In order to make the model as widely applicable as possible, we need to make as few restrictions as possible on functions h_v and g_v in model (4). In order to facilitate analysis, we need functions h_v and g_v to have the following Properties 1-4, which are both intuitive and natural.

Property 1. For $\forall v$, $h_v(i, \beta_v(t))$ and $g_v(i, \alpha_v(t), \Gamma(t))$ have continuous first and second derivatives with respect to i .

Property 1 is the baseline for analytical treatment.

Property 2. Consider $\forall v$ and $\forall u \neq v$, (i) $g_v(\mathbf{0}, \alpha_v(t), \Gamma(t)) = \alpha_v(t)$, $\partial g_v(i(t), \alpha_v(t), \Gamma(t)) / \partial i_u = 0$ when $\gamma_{vu}(t) = 0$, and $\partial h_v(i(t), \beta_v(t)) / \partial i_u = 0$; (ii) for any $\delta > 0$, $\partial g_v(i(t), \alpha_v(t), \Gamma(t)) / \partial i_u > \delta_1$ holds for some $\delta_1 > 0$ when $\gamma_{vu}(t) > \delta$.

Part (i) of Property 2 abstracts the intuition that the probability node v gets compromised at time t is independent of the state of u when u cannot attack node v or $(u, v) \notin E(t)$, and the probability node u becomes secure at time t is independent of the states of the other nodes at time t . Part (ii) of Property 2 abstracts the intuition that the probability v becomes compromised at time t increases with the probability u is compromised when u can attack v or $(u, v) \in E(t)$.

Property 3. Consider $\forall v$, $g_v(i, \alpha_v(t), \Gamma(t))$ is subhomogeneous with respect to i and $D_i(g_v(i, \beta_v(t)) + h_v(i, \alpha_v(t), \Gamma(t))) = [\partial(g_v(i, \beta_v) + h_v(i, \alpha_v, \Gamma)) / \partial i_u]_{u \in V} > 0$.

Property 3 abstracts the intuition that there is always a nonzero probability for a compromised node v to become secure because reactive defenses have a nonzero probability to succeed or $\beta_v(t) \geq 0$.

In order to prove the global attractivity of model (4) with time-dependent parameters, we need parameters $y(t)$ to satisfy the following property:

Property 4 (ergodicity of $\{y(t)\}_{t \geq 0}$). $\{y(t)\}_{t \geq 0}$ is ergodic, where $y(t) = [\alpha(t), \beta(t), \text{vec}(A(t))^\top, \text{vec}(\Gamma(t))^\top]$.

Note that Properties 1-3 are naturally satisfied by the \prod -model and the \sum -model with time-dependent parameters, namely systems or models (6). This is because Properties 1-3 naturally extend their time-independent counterparts, which are satisfied by the \prod -model and the \sum -model with time-independent parameters. We will show that Property 4 (i.e., ergodicity) is close to, if not, the necessary condition for the global attractivity of model (4), by presenting a numerical example to show that its violation disrupts the global attractivity.

The validation of these four properties, while intuitive, is an orthogonal research problem to the present characterization study and will be investigated in the future.

C. Model (4) Is Strongly Subhomogeneous

We start by showing that model (4) is strongly subhomogeneous under Properties 1 and 3.

Theorem 2. Model (4) under Properties 1 and 3 is strongly subhomogeneous in $[0, 1]^n$.

The proof of Theorem 2 is similar to the proof of Lemma 4 in [13], which considers the unified model with time-independent parameters. This is because the proof does not need to make any restrictions on the time-dependence of the parameters, meaning that the proof is equally applicable to both time-independent parameters (the setting of [13]) and our setting of time-dependent parameters.

D. Special equilibrium $\mathbf{0}$ Is Globally Stable

Theorem 3 below shows that when there are no pull-based attacks, namely $\alpha_v(t) = 0, \forall v \in V, \forall t$, the special equilibrium $\mathbf{0}$ of model (4) is globally stable under a certain condition. Note that $\mathbf{0}$ is no equilibrium when some nodes are subject to pull-based attacks, namely that the average of $\alpha_v(t)$ over t is positive for some $v \in V$. Proof of Theorem 3 is deferred to Appendix B.

Theorem 3 (equilibrium $\mathbf{0}$ is globally stable under Properties 1-3). Consider model (4) and $\alpha_v(t) = 0, \forall v \in V, \forall t$. For any $y \in Y$ with $\mu(D_i f(\mathbf{0}, y)) < 0$, equilibrium $\mathbf{0}$ is globally stable in $[0, 1]^n$, namely that every trajectory $\psi(t, i_0, y)$ of model (4) converges to $\mathbf{0}$ regardless of the initial value $i(0) \in [0, 1]^n$.

E. Model (4) Is Globally Attractive

Theorem 4 (main result: the unified model with time-dependent parameters is globally attractive under Properties 1-4). Model (4) is almost surely globally attractive in $[0, 1]^n \setminus \{\mathbf{0}\}$, meaning that for almost every $y \in Y$, there exists a unique trajectory $\psi(t, i_y^*, y) \in [0, 1]^n$ such that $\lim_{t \rightarrow \infty} \|\psi(t, i_0, y) - \psi(t, i_y^*, y)\| = 0$ holds for any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$. Moreover, $\forall v \in V$, $[\psi(t, i_y^*, y)]_v \neq 0$ as long as node v is subject to pull-based attacks, meaning that the average of $\alpha_v(t)$ over time is positive.

Note that equilibrium $\mathbf{0}$ can be seen as a special case of a trajectory, but we separate its treatment (Theorem 3) from the treatment of the global attractivity in $[0, 1]^n \setminus \{\mathbf{0}\}$ (Theorem 4) because the former, as a special case, can be proven without requiring the parameters $\{y(t)\}_{t \geq 0}$ to be ergodic (Property 4).

The proof of Theorem 4 is deferred to Appendix G. In order to prove Theorem 4, we need the following Lemmas 2-5, whose proofs are respectively deferred to Appendices C-F. In the following lemmas, we let $M(\Gamma)$ denote the mean value of the ergodic process $\{\Gamma(t)\}_{t \geq 0}$ and call the weighted graph $\mathcal{G}(M(\Gamma))$ the mean attack-defense structure of $(G(t), \Gamma(t))$. Lemmas 2-5 cope with different settings of the parameters (following the ‘‘divide and conquer’’ strategy):

there are pushed-based and pull-based attacks vs. there are only pushed-based attacks (i.e., no nodes are subject to pull-based attacks); the mean attack-defense structure is strongly connected vs. it is not strongly connected. Lemmas 2-3 deal with the case that there are no pull-based attacks, namely $\alpha_v(t) = 0, \forall v \in V, \forall t$; Lemma 4 deals with the case that there is at least some node $v \in V$ that is subject to pull-based attack, namely that there is at least one $v \in V$ such that the average of $\alpha_v(t)$'s over time t is positive; Lemma 5 deals with the case that the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ has two strongly connected components, where the mean is averaged over time. Note that it suffices to consider two strongly connected components because multiple strongly connected components can be treated in the same fashion and any attack-defense structure can be partitioned into strongly connected components.

1) *Global attractivity when there are no pull-based attacks:* Theorem 1 says that model (4) is globally attractive when the dynamics is *uniformly persistent* (Definition 5) and satisfies conditions (I) and (II) of Theorem 1. We first prove that model (4) is *uniformly persistent* when $\mu(D_i f(\mathbf{0}, y)) > 0$ and the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ is strongly connected.

Lemma 2 (model (4) is uniformly persistent when there are no pull-based attacks, under Properties 1-4). *If $\mu(D_i f(\mathbf{0}, y)) > 0$ and the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ is strongly connected, trajectories of model (4) with nonzero initial values are uniformly persistent, namely $\liminf_{t \rightarrow +\infty} \psi(t, i_0, y) \gg \mathbf{0}, \forall i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$ and $\forall y \in Y$.*

Lemma 3 below shows that model (4) is *globally attractive* (Definition 6) when Properties 1-4 hold and the mean attack-defense structure is strongly connected.

Lemma 3 (model (4) is globally attractive when there are no pull-based attacks under Properties 1-4). *Suppose the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ is strongly connected.*

- (i) *If $\mu(D_i f(\mathbf{0}, y)) > 0$, then there exists, for almost every $y \in Y$, a positive trajectory $\psi(t, i_y^*, y) \in (0, 1]^n$ that is globally attractive in $[0, 1]^n \setminus \{\mathbf{0}\}$.*
- (ii) *If $\mu(D_i f(\mathbf{0}, y)) = 0$, then there exists, for almost every $y \in Y$, a trajectory $\psi(t, i_y^*, y) \in [0, 1]^n$ that is globally attractive in $[0, 1]^n \setminus \{\mathbf{0}\}$. Moreover, if $\psi(t, i_{y_0}^*, y_0) \in (0, 1]^n$ for some $y_0 \in Y$, then for almost every $y \in Y$, its associated globally attractive trajectory $\psi(t, i_y^*, y) \in (0, 1]^n$.*
- (iii) *If $\mu(D_i f(\mathbf{0}, y)) < 0$, the origin $\mathbf{0}$ is almost surely globally attractive, i.e., $\lim_{t \rightarrow \infty} \psi(t, i_0, y) = \mathbf{0}$ holds for any $i_0 \in [0, 1]^n$ and almost every $y \in Y$.*

2) *Global attractivity when there are pull-based attacks:* From the ergodicity and non-negativity of the $\alpha_v(t)$'s, it follows that the mean value of $\{\alpha_v(t)\}_{t \geq 0}$ over t , denoted by $M(\alpha_v)$, is zero, namely $M(\alpha_v) = 0$, if and only if $\alpha_v(t) = 0$ almost surely. Therefore, when $\alpha_v(t)$ is ergodic, node v is subject to pull-based attacks if and only if the mean value $M(\alpha_v) > 0$. In what follows, we will use $\alpha_v(t) > 0$ to denote that the mean value $M(\alpha_v)$ is positive. Let $V_{\alpha > 0} = \{v : M(\alpha_v) > 0\}$ denote the set of nodes that are subject to pull-based attacks. The global attractivity of model

(4) when $V_{\alpha > 0} \neq \emptyset$ is established by the following lemma, while noting that $\mathbf{0}$ is no longer an equilibrium of model (4) when some nodes are subject to pull-based attacks.

Lemma 4 (model (4) is globally attractive when there are pull-based attacks under Properties 1-4). *If $V_{\alpha > 0} \neq \emptyset$ and the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ is strongly connected, then there exists, for almost every $y \in Y$, a unique positive trajectory $\psi(t, i_y^*, y)$ that is globally attractive in $[0, 1]^n$, namely $\lim_{t \rightarrow \infty} \|\psi(t, i_0, y) - \psi(t, i_y^*, y)\| = 0$ holds $\forall i_0 \in [0, 1]^n$.*

3) *Global attractivity when the mean attack-defense structure is not strongly connected:* Lemmas 3-4 show model (4) is globally attractive when the mean attack-defense structure $\mathcal{G}(M(\Gamma))$ is strongly connected. Now we consider the case that $\mathcal{G}(M(\Gamma))$ is not strongly connected, but can be divided into two strongly connected components (SCCs), denoted by SCC_1 and SCC_2 . When there are no links between SCC_1 and SCC_2 , the global attractivity for each $SCC_k, k = 1, 2$ is obtained from Lemma 3 and Lemma 4 directly by treating each SCC as an attack-defense structure. Therefore, we only need to consider the case that there exist links between these two strongly connected components of $\mathcal{G}(M(\Gamma))$.

From the Perron-Frobenius theory [44], $M(\Gamma)$ can be written in the lower-triangular block form,

$$M(\Gamma) = P \begin{bmatrix} \tilde{\Gamma}_{11} & 0 \\ \tilde{\Gamma}_{21} & \tilde{\Gamma}_{22} \end{bmatrix} P^\top$$

where P is a permutation matrix, $\tilde{\Gamma}_{kk}$ is irreducible and corresponds to SCC_k for $k = 1, 2$, and $\tilde{\Gamma}_{21} \neq 0$. It follows from Property 2 that for any $v \neq u, \partial f_v(\mathbf{0}, y)/\partial i_u = \partial g_v(\mathbf{0}, \alpha_v, \Gamma)/\partial i_u = 0$ holds if $\gamma_{vu}(t) = 0$, implying that the Jacobian matrix $D_i f(\mathbf{0}, y)$ can be written as

$$D_i f(\mathbf{0}, y) = P \begin{bmatrix} Df_{11} & 0 \\ Df_{21} & Df_{22} \end{bmatrix} P^\top,$$

where Df_{kk} corresponds to SCC_k for $k = 1, 2$. Let \mathcal{V}_{SCC_k} denote the set of nodes in SCC_k and $|SCC_k|$ denote the number of nodes in SCC_k , where $k = 1, 2$. The following lemma considers the general case of $\alpha_v(t) \geq 0, \forall v \in V$.

Lemma 5 (model (4) is globally attractive when the mean attack-defense structure has two strongly connected components, under Properties 1-4). *Suppose $\mathcal{G}(M(\Gamma))$ consists of two strongly connected components SCC_1 and SCC_2 such that (without loss of generality) at least one node in SCC_1 has a path to a node in SCC_2 . For SCC_1 , there exists a globally attractive trajectory $\phi(t, i_{SCC_1}^*, y) \in [0, 1]^{|SCC_1|}$ such that*

- (i) *if $\phi(t, i_{SCC_1}^*, y) = \mathbf{0}_{|SCC_1|}, \mathcal{V}_{SCC_2} \cap V_{\alpha > 0} = \emptyset$ and $\mu(Df_{22}) < 0$, the origin $\mathbf{0}_n$ is globally attractive for $\mathcal{V}_{SCC_1} \cup \mathcal{V}_{SCC_2}$, meaning that every trajectory of model (4) converges to $\mathbf{0}_n$ regardless of the initial values;*
- (ii) *otherwise, there is a trajectory $\psi(t, i^*, y)$ that is globally attractive in $[0, 1]^n \setminus \{\mathbf{0}_n\}$, where $[\psi(t, i^*, y)]_{SCC_1} = \phi(t, i_{SCC_1}^*, y)$. Moreover, if $\mathcal{V}_{SCC_2} \cap V_{\alpha > 0} \neq \emptyset, [\psi(t, i^*, y)]_{SCC_2} \in (0, 1]^{|SCC_2|}$ is globally attractive in $[0, 1]^{|SCC_2|}$.*

F. Stronger Assumptions Leading to Stronger Results

Now we show that stronger results can be obtained when the parameters exhibit stronger properties than ergodicity, such as (almost) periodic (Definition 2).

Theorem 5 (model (4) is globally attractive and almost periodic when its parameters are almost periodic). *Consider model (4) under Properties 1-3 and $y(t)$ is almost periodic (i.e., a property stronger than Property 4): (i) there exists an almost periodic trajectory $\psi(t, i^*, y) \in [0, 1]^n$ that is globally attractive in $[0, 1]^n \setminus \{\mathbf{0}\}$, namely that $\lim_{t \rightarrow \infty} \|\psi(t, i_0, y) - \psi(t, i^*, y)\| = 0$ holds for any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$; (ii) for any $\epsilon > 0$, there exists $M > 0$ so that the ϵ -translation numbers of $y(t)$ are $M\epsilon$ -translation numbers of the globally attractive trajectory $\psi(t, i^*, y)$.*

Proof of Theorem 5 is deferred to Appendix H. Since periodic functions are almost periodic by setting $\epsilon = 0$ in Definition 2, we obtain the following corollary.

Corollary 1 (model (4) is globally attractive and periodic when its parameters are periodic). *Consider model (4) under Properties 1-3. Suppose $y(t)$ is periodic with period w , then there exists an w -periodic trajectory $\psi(t, i^*, y)$ of model (4) that is globally attractive in $[0, 1]^n \setminus \{\mathbf{0}\}$, namely that $\lim_{t \rightarrow \infty} \|\psi(t, i_0, y) - \psi(t, i^*, y)\| = 0$ holds for any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$.*

G. Bounding the Globally Attractive Positive Trajectory

When given all of the parameters, one can numerically compute the globally attractive trajectory. In practice, we may not know the values of all parameters (i.e., the matter of partial vs. full information). In this case, a useful alternative is to bound the globally attractive trajectory because, for example, we can treat the upper bound as the worst case scenario in cyber defense decision-making. In order to derive such bounds, we need h_v and g_v to have the following intuitive property:

Property 5 (properties of h_v and g_v needed for bounding the globally attractive trajectory). *Consider $g_v(i, \alpha_v, \Gamma)$ and $h_v(i, \beta_v)$, it holds that*

- (i) $\partial g_v / \partial \alpha_v \geq 0$ and $\partial g_v / \partial \gamma_{vu} \geq 0$ for $\forall u, v \in V$, which reflects the intuition that the probability node v getting compromised increases with the pull-based and push-based attack capabilities; and
- (ii) $\partial h_v / \partial \beta_v \geq 0$ for $\forall v \in V$, which reflects the intuition that when everything else is fixed, the probability a compromised node v becoming secure increases with the reactive defense capability.

To simplify the presentation, in the rest of this subsection

we will use the following notations. For any $u, v \in V$, we set

$$\begin{aligned} \underline{\alpha}_v &= \min_{t \geq 0} \alpha_v(t), & \bar{\alpha}_v &= \max_{t \geq 0} \alpha_v(t), \\ \underline{\beta}_v &= \min_{t \geq 0} \beta_v(t), & \bar{\beta}_v &= \max_{t \geq 0} \beta_v(t), \\ \underline{\gamma}_{vu} &= \min_{t \geq 0} \gamma_{vu}(t), & \bar{\gamma}_{vu} &= \max_{t \geq 0} \gamma_{vu}(t), \\ \underline{\Gamma} &= [\underline{\gamma}_{vu}]_{v, u \in V}, & \bar{\Gamma} &= [\bar{\gamma}_{vu}]_{v, u \in V}, \\ i_v^{\min} &= \inf_{t \geq 0} \{i_v(t)\}, & i_v^{\max} &= \sup_{t \geq 0} \{i_v(t)\}, \\ i_{\min} &= [i_1^{\min}, \dots, i_n^{\min}], & i_{\max} &= [i_1^{\max}, \dots, i_n^{\max}], \\ \underline{h}_v(\beta_v) &= \min_{i \in [i_{\min}, i_{\max}]} h_v(i, \beta_v), & \bar{h}_v(\beta_v) &= \max_{i \in [i_{\min}, i_{\max}]} h_v(i, \beta_v). \end{aligned}$$

The basic idea is to leverage i_{\min} and i_{\max} to derive the bounds. When i_{\min} and i_{\max} are unknown, we can set $i_{\min} = \mathbf{0}$ and $i_{\max} = \mathbf{1}$. On the other hand, the properties $\partial g_v / \partial i_u \geq 0$, $\partial g_v / \partial \alpha_v \geq 0$ and $\partial g_v / \partial \gamma_{vu} \geq 0$ lead to that g_v is monotone with respect to i , α_v and γ_{vu} , which implies

$$g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma}) \leq g_v(i(t), \alpha_v, \Gamma) \leq g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma}).$$

Then, it follows from $\partial h_v / \partial \beta_v \geq 0$ that

$$\underline{h}_v(\underline{\beta}_v) \leq h_v(i, \beta_v) \leq \bar{h}_v(\bar{\beta}_v).$$

The following theorem can bound the dynamics when only knowing the lower and upper bounds of parameters $\alpha_v(t)$, $\beta_v(t)$, and $\gamma_{vu}(t)$, $\forall u, v \in V, \forall t$, while the $G(t)$'s are given. Proof of Theorem 6 is deferred to Appendix I.

Theorem 6 (bounding the globally attractive trajectory). *Let $i_v(t)$ be the solution to model (4) and $\bar{i}_v(t)$ and $\underline{i}_v(t)$ be the upper and lower bounds of $i_v(t)$. Then under Property 5, we have $\underline{i}_v(t) \leq i_v(t) \leq \bar{i}_v(t)$ for $\forall v \in V, t \in \mathbb{R}_{\geq 0}$, where*

$$\underline{i}_v(t) = \exp(-\underline{\mathcal{A}}_v t) \left[\underline{i}_v(0) - \frac{\underline{\mathcal{B}}_v}{\underline{\mathcal{A}}_v} \right] + \frac{\underline{\mathcal{B}}_v}{\underline{\mathcal{A}}_v} \quad (7)$$

$$\bar{i}_v(t) = \exp(-\bar{\mathcal{A}}_v t) \left[\bar{i}_v(0) - \frac{\bar{\mathcal{B}}_v}{\bar{\mathcal{A}}_v} \right] + \frac{\bar{\mathcal{B}}_v}{\bar{\mathcal{A}}_v} \quad (8)$$

with $\underline{\mathcal{A}}_v$, $\bar{\mathcal{A}}_v$, $\underline{\mathcal{B}}_v$, and $\bar{\mathcal{B}}_v$ satisfying

$$\begin{cases} \underline{\mathcal{A}}_v = \bar{h}_v(\bar{\beta}_v) + g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma}), \\ \underline{\mathcal{B}}_v = g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma}), \\ \bar{\mathcal{A}}_v = \underline{h}_v(\underline{\beta}_v) + g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma}), \\ \bar{\mathcal{B}}_v = g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma}). \end{cases}$$

H. Relationship between the Literature Results and Ours

Now we show that our results supersede the state-of-the-art results because they are equivalent to corollaries of our results.

Corollary 2 (corollary of our Theorem 4 and Lemmas 3-5 is equivalent to Theorem 3 in [13]). *Consider model (4) under Properties 1-3. Suppose parameters $\alpha_v(t)$, $\beta_v(t)$, $\gamma_{vu}(t)$ are time-independent for all $u, v \in V$. Suppose the attack-defense structure $G(t)$ is also time-independent and contains K strongly connected components, denoted by SCC_k for $k = 1, \dots, K$, and the Jacobian matrix $D_i f(\mathbf{0}, y)$ has the following Perron-Frobenius form*

$$D_i f(\mathbf{0}, y) = P \begin{bmatrix} Df_{11} & 0 & \dots & 0 \\ Df_{21} & Df_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Df_{K1} & Df_{K2} & \dots & Df_{KK} \end{bmatrix} P^\top \quad (9)$$

where P is a permutation matrix, and Df_{kk} corresponds to SCC_k for $k = 1, \dots, K$. Let R_k be the indices of the SCC 's that have links pointing to SCC_k . Then, for each SCC_k , we have:

- (i) $\mathbf{0}_{|SCC_k|}$ is globally asymptotically stable in $[0, 1]^n$ if one of the following conditions holds:
 - $\mathcal{V}_{SCC_k} \cap \mathcal{V}_{\alpha > 0} = \emptyset$, $\mu(Df_{kk}) \leq 0$ and $SCC_{R_k} = \emptyset$;
 - $\mathcal{V}_{SCC_k} \cap \mathcal{V}_{\alpha > 0} = \emptyset$, $\mu(Df_{kk}) \leq 0$, $SCC_{R_k} \neq \emptyset$ and $\mathbf{0}_{|SCC_r|}$ is globally asymptotically stable for every $r \in R_k$.
- (ii) SCC_k has a unique positive equilibrium that is globally asymptotically stable in $[0, 1]^{|SCC_k|}$ if $\mathcal{V}_{SCC_k} \cap \mathcal{V}_{\alpha > 0} \neq \emptyset$.
- (iii) SCC_k has a unique positive equilibrium that is globally asymptotically stable in $[0, 1]^{|SCC_k| \setminus \{0\}}$ if one of the following conditions holds:
 - $SCC_k \cap \mathcal{V}_{\alpha > 0} = \emptyset$ and $\mu(Df_{kk}) > 0$;
 - $SCC_k \cap \mathcal{V}_{\alpha > 0} = \emptyset$, $SCC_{R_k} \neq \emptyset$ and $\mathbf{0}_{SCC_r}$ is not globally asymptotically stable for every $r \in R_k$.

The state-of-the-art result of the Σ -model with time-dependent $G(t)$ is given in [30, 31], and the state-of-the-art result of the Π -model with time-dependent $G(t)$ is given in [18, 27, 28, 29]. However, these studies only investigate the stability of the equilibrium $\mathbf{0}$. Among these studies, [27, 28, 29, 30] consider the discrete-time model and show that the equilibrium $\mathbf{0}$ is stable if the *joint spectral radius* of the set of system matrices is smaller than 1. For a discrete-time system $z(t+1) = C(t)z(t)$, $t \in \mathbb{N}$, let $\mu_d(C(t)) = \lim_{n \rightarrow \infty} 1/n \log(\|C(n) \cdots C(1)C(0)\|)$ be the associated maximum Lyapunov exponent (MLE) and $\rho(C(t))$ be the joint spectral radius given in [27, 28, 29, 30]. In fact, the joint spectral radius can be seen as the discrete-time version of the MLE; i.e., $\rho(C(t)) < 1$ is equivalent to $\mu_d(C(t)) < 0$ and thus their result is a special case of our Theorem 3.

Corollary 3 (corollary of our Theorem 3 is equivalent to Theorem 1 in [27], Theorem 2 in [28], Theorem 1 in [29] and Theorem 2.1 in [30]). *Consider the discrete-time version of model (4). Suppose $\alpha_v(t) = 0$ and $\gamma_{vu}(t) = \gamma(t)$, $\forall u, v \in V$, $\forall t$. Let $B(t) = \text{diag}(\beta_1(t), \dots, \beta_n(t))$. If $\rho((I - B(t)) + \gamma(t)A(t)) < 1$, the dynamics of the discrete-time model globally converges to equilibrium $\mathbf{0}$.*

The Σ -model with time-dependent parameters is studied in [31] and shown to converge to the equilibrium $\mathbf{0}$ when the time-dependent parameters satisfy some specific conditions. For symmetric matrices $C(t)$, $t \geq 0$, we have $\mu(C(t)) \leq \sup_{t \geq 0} \lambda_1(C(t))$, meaning that the following corollary of our Theorem 3 is equivalent to the result of [31].

Corollary 4 (corollary of our Theorem 3 is equivalent to Theorem 1 in [31]). *Consider model (4) while instantiating its term $g_v(i(t), \alpha_v(t), \Gamma(t))$ as the Σ -model with time-dependent parameters. Suppose $\alpha_v(t) = 0$ and $\Gamma(t)$ being symmetric for $\forall v \in V$, $\forall t$. Let $B(t) = \text{diag}(\beta_1(t), \dots, \beta_n(t))$. If $\sup_{t \geq 0} \lambda_1(\Gamma(t) - B(t)) < 0$, the dynamics of (4) globally converges to equilibrium $\mathbf{0}$.*

The Π -model with time-dependent parameters is considered in [18], which shows when the dynamics converges to the

equilibrium $\mathbf{0}$. The following corollary of our Theorems 3 and 4. is equivalent to the result of [18].

Corollary 5 (corollary of our Theorems 3 and 4 is equivalent to Theorem 1 in [18]). *Consider model (4) while instantiating its term $g_v(i(t), \alpha_v(t), \Gamma(t))$ as the Π -model with time-dependent parameters. Suppose $\alpha_v(t) = 0$ and $\gamma_{vu}(t) = \gamma(t)$, $\forall u, v \in V$, $\forall t$. Let $B(t) = \text{diag}(\beta_1(t), \dots, \beta_n(t))$. If $\mu(\gamma(t)A(t) - B(t)) < 0$, model (4) in this special case converges to equilibrium $\mathbf{0}$; if $\mu(\gamma(t)A(t) - B(t)) > 0$ and the linear system $dz(t)/dt = (\gamma(t)A(t) - B(t))z(t)$ is ergodic, equilibrium $\mathbf{0}$ is unstable.*

I. Systematizing Knowledge

We use Figure 4 to systematize the relationship between the properties, lemmas, theorems, and corollaries (i.e., their equivalent literature results).

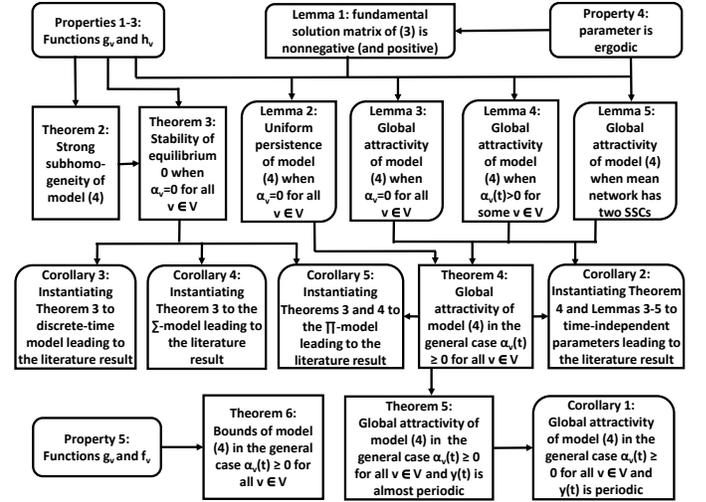


Fig. 4: Relationship between the mathematical properties, lemmas, theorem, and corollaries presented in the paper.

IV. NUMERICAL EXAMPLES

We use numerical results to confirm our analytic results. In our experiments, we use the Euler method for the numerical simulation of model (4) by setting the iteration step as 0.05. In order to succinctly present the experimental result, we plot the dynamics of $\langle i(t) \rangle = \sum_v i_v(t)/|V|$, which is the fraction of compromised nodes at time t . In our experiment, we set the initial fraction of compromised nodes as $\langle i(0) \rangle \in \{0.25, 0.5, 0.75\}$, meaning that 25%, 50%, and 75% randomly chosen nodes are in the compromised state at time $t = 0$.

A. Confirming Global Attractivity

In our experiments, we use the Gnutella05 peer-to-peer network <http://snap.stanford.edu/data/> as the initial attack-defense structure $G(0)$, where $|V| = 8846$ nodes and $|E(0)| = 31839$ arcs. In order to generate $G(t)$ for $t > 0$, we randomly add or delete 2% of the arcs of the attack-defense structure after every 10 simulation time units. In our experiments, we consider the Σ -model with time-dependent parameters $h_v(i, \beta_v) = \beta_v(t)$

and $g_v(i, \alpha_v, \Gamma)$ given in model (6), which satisfies Properties 1-3 as required by the main results (i.e., Theorem 4).

In the first experiment, we set $\alpha_v(t) = 0, \forall v, \forall t$, and use the following parameter sets

- (p1): $\beta_v(t) = 0.1 \sin(t) + 0.1 \sin(\sqrt{2}t) + 0.5, \forall v \in V$, and $\gamma_{uv}(t) = 0.05 \sin(\pi t/5) + 0.1, \forall (u, v) \in E(t)$;
- (p2): $\beta_v(t) = 0.1 \sin(t) + 0.1 \sin(\sqrt{2}t) + 0.4, \forall v \in V$, and $\gamma_{uv}(t) = 0.05 \sin(\pi t/5) + 0.1, \forall (u, v) \in E(t)$;
- (p3): $\beta_v(t) = 0.1 \sin(t) + 0.1 \sin(\sqrt{2}t) + 0.1, \forall v \in V$, and $\gamma_{uv}(t) = 0.05 \sin(\pi t/5) + 0.1, \forall (u, v) \in E(t)$;
- (p4): $\gamma_{uv}(t) = 0.3 \sin(\pi t/5) + 0.7, \forall u, v$; multiple $\beta_v(t)$'s that will be specified below.

Note that parameter sets (p1), (p2) and (p3) correspond to the cases $\mu(D_i f(\mathbf{0}, y(t))) < 0$, $\mu(D_i f(\mathbf{0}, y(t))) = 0$ and $\mu(D_i f(\mathbf{0}, y(t))) > 0$, respectively.

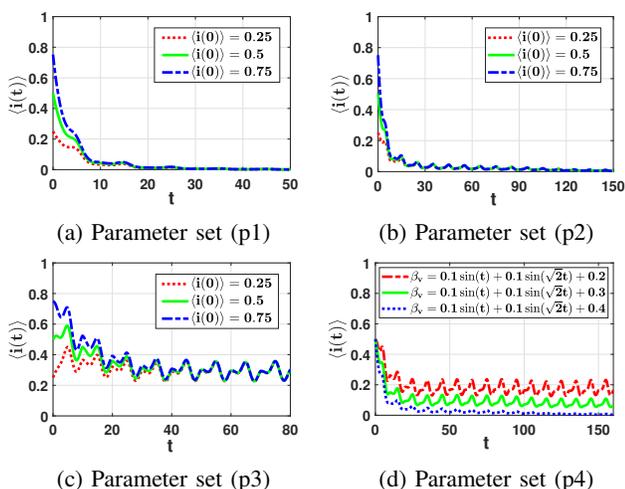


Fig. 5: Global attractivity of the \sum -model (6) with time-dependent parameters except that there are no pull-based attacks, namely $\alpha_v(t) = 0, \forall v \in V, \forall t$.

Figures 5(a), 5(b) and 5(c) plot the dynamics of the fraction of compromised nodes over time, namely $\langle i(t) \rangle$, with parameter sets (p1), (p2), (p3) and different initial infection values. We observe that the the dynamics always converges to a unique trajectory. It reinforces the result that the dynamics converges to the equilibrium $\mathbf{0}$ when $\mu(D_i f(\mathbf{0}, y(t))) < 0$, attracts to a positive trajectory when $\mu(D_i f(\mathbf{0}, y(t))) > 0$, and attracts to a trajectory (possibly an equilibrium $\mathbf{0}$ or non-negative trajectory) when $\mu(D_i f(\mathbf{0}, y(t))) = 0$. Figure 5(d) plots the dynamics of $\langle i(t) \rangle$ with multiple $\beta_v(t)$'s, and shows that the globally attractive trajectory converges to $\mathbf{0}$ as $\beta_v(t)$ increases. This reinforces the intuition that a larger $\beta_v(t)$ leads to a smaller $\mu(D_i f(\mathbf{0}, y(t)))$ and $\langle i(t) \rangle$ as well as the convergence to the equilibrium $\mathbf{0}$.

In the second experiment, we set $\alpha_v(t) = 0.1 \sin(3t) + 0.1 \sin(\sqrt{3}t) + 0.2$ if $v \in V_{\alpha>0}$, set $\beta_v(t)$ and $\gamma_{vu}(t)$ as in the aforementioned parameter set (p2), and consider $|V_{\alpha>0}|/|V| \in \{0.25, 0.5, 0.75\}$, namely that 25%, 50%, and 75% randomly chosen nodes are subject to pull-based attacks, respectively. Figure 6(a) plots the dynamics of $\langle i(t) \rangle$ with $|V_{\alpha>0}|/|V| = 0.5$ and different initial values, and shows that the dynamics converges to a unique positive trajectory. Figure

6(b) plots the dynamics of $\langle i(t) \rangle$ under different pull-based attack capabilities, and shows a positive correlation between the fraction of the compromised nodes and the pull-based attack capabilities.

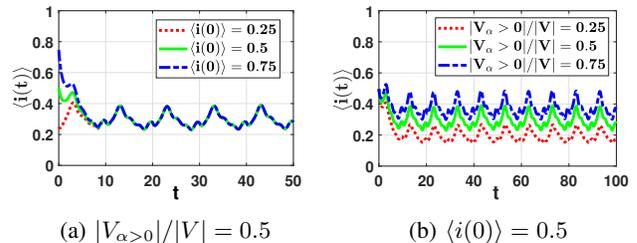


Fig. 6: Dynamics of the \sum -model (6) with time-dependent parameters, including different pull-based attack capabilities.

B. Confirming Bounds

In order to characterize the impact of time-dependent functions and pull-based attacks on the tightness of the bounds, we consider the following four parameter sets (p5)-(p8):

- (p5): $\alpha_v(t) = 0.05 \sin(3t) + 0.05 \sin(\sqrt{3}t) + 0.5, \forall v \in V_{\alpha>0}$ where $|V_{\alpha>0}|/|V| = 0.5$; $\beta_v(t) = 0.05 \sin(t) + 0.05 \sin(\sqrt{2}t) + 0.5, \forall v \in V$; $\gamma_{uv}(t) = 0.1 \sin(\pi t/5) + 0.5, \forall (u, v) \in E(t)$;
- (p6): $\alpha_v = 0, \beta_v(t) = 0.1 \sin(t) + 0.1 \sin(\sqrt{2}t) + 0.4, \forall v \in V$; $\gamma_{uv}(t) = 0.1 \sin(\pi t/5) + 0.5, \forall (u, v) \in E(t)$;
- (p7): $\{\alpha_v(t)\}_{t \geq 0} \sim U([0.1, 0.3]), \forall v \in V_{\alpha>0}$ where $|V_{\alpha>0}|/|V| = 0.5$; $\{\beta_v(t)\}_{t \geq 0} \sim U([0.4, 0.7]), \forall v \in V$; $\gamma_{uv}(t) = 0.1, \forall (u, v) \in E(t)$;
- (p8): $\alpha_v = 0, \{\beta_v(t)\}_{t \geq 0} \sim U([0.4, 0.7]), \forall v \in V$; $\gamma_{uv}(t) = 0.1, \forall (u, v) \in E(t)$.

Note that in these settings, it holds that $\{\eta(t)\}_{t \geq 0} \sim U([a, b])$ if $\eta(t)$ satisfies (i) $\eta(t) = \eta_k$ when $t \in [k, k+1), k \in \mathbb{N}$ and (ii) $\eta_k, k \in \mathbb{N}$ follow the uniformly distribution in $[a, b]$ independently. Moreover, parameter sets (p7) and (p8) satisfy the ergodic property.

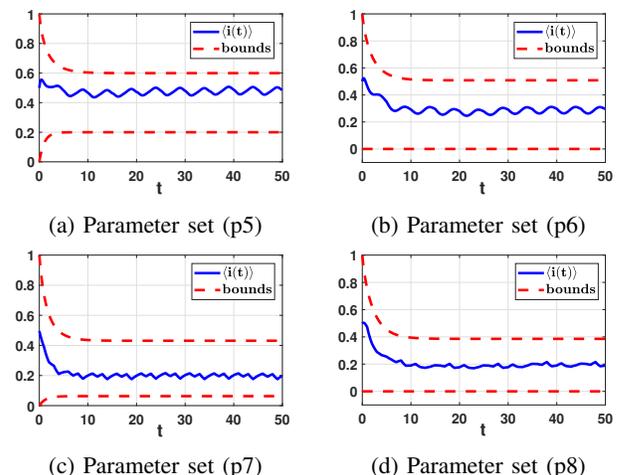


Fig. 7: Dynamics of the \sum -model (6) with time-dependent parameters and their bounds given by Theorem 6.

Figure 7 plots the dynamics of $\langle i(t) \rangle$ with parameter sets (p5)-(p8) and the corresponding bounds given by Theorem 6. The lower bounds in Figure 7 (b) and (d) are 0 because $i_{\min} = 0$ and $\alpha_v(t) = 0$ in parameter sets (p6) and (p8). We observe that the bounds in Figure 7 can be very loose, perhaps because $i_{\max} = 1$ and $i_{\min} = 0$ in this example, which reflects the lack of information on the globally attractive trajectory.

C. Are the Sufficient Conditions Necessary?

It is known [16] that some cybersecurity dynamics can exhibit bifurcation and chaos. Since we leverage subhomogeneity (Property 3) and ergodicity (Property 4) to obtain the global attractivity result for the unified dynamics with time-dependent parameters, it makes us wonder how far these sufficient conditions are from being necessary. In what follows we use numerical examples to show that violating these properties can cause violation of global attractivity, hinting subhomogeneity and ergodicity may be necessary; the rigorous treatment of this is a difficult task and left for future research.

For constructing examples, it suffices to consider a special kind of time-independent attack-defense structures G in Erdős-Rényi (ER) random graph. Specifically, we consider an ER structure with $n = 1,000$ nodes and edge probability $p = 0.1$. The ER graph is then interpreted as a directed graph. We consider the Σ -model as an example.

First, we empirically show that subhomogeneity (required by Property 3) may be necessary for global attractivity. Let us consider the following functions for the Σ -model:

$$h_v(i, \beta_v) = \beta_v(t),$$

$$g_v(i, \alpha_v, \Gamma) = \alpha_v(t) + \left(\frac{\sum_{u \in N_v(t)} \gamma_{vu}(t) a_{vu}(t) i_u(t)}{\max\{|N_v(t)|, 1\}} \right)^2.$$

Note that the preceding g_v is *not* subhomogeneous. We consider the following two combinations:

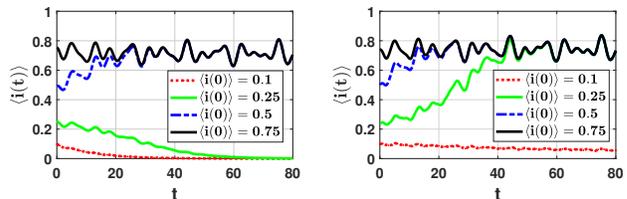
- (p9): $\alpha_v = 0, \forall v; \beta_v(t) = 0.05 \sin(t) + 0.05 \sin(\sqrt{2}t) + 0.1, \forall v \in V; \gamma_{uv}(t) = 0.1 \sin(\pi t/5) + 0.7, \forall (u, v) \in E(t);$
- (p10): $\alpha_v(t) = 0.1 \sin(3t) + 0.1 \sin(\sqrt{3}t) + 0.1, \forall v \in V_{\alpha>0}$ where $|V_{\alpha>0}|/|V| = 0.2$; the other parameters are the same as in (p9).

Figure 8(a) plots the dynamics of $\langle i(t) \rangle$ with parameter set (p9) but different initial values $\langle i(0) \rangle$. We observe that the dynamics is not globally attractive because different initial values lead to different trajectories. Figure 8(b) plots the dynamics of $\langle i(t) \rangle$ with parameter set (p10) but different initial values $\langle i(0) \rangle$. We observe that the dynamics is not globally attractive because different initial values lead to different trajectories. These experiments hint that subhomogeneity may be necessary for global attractivity.

Second, we empirically show that ergodicity may be necessary for global attractivity. Let us consider the Σ -model (6) without pull-based attacks, namely

- (p11): $\alpha_v(t) = 0, \forall v \in V; \gamma_{uv}(t) = 0.3, \forall (u, v) \in E(t)$ and $t \geq 0$;

$$\{\beta_v(t)\}_{t \geq 0} \sim \begin{cases} U([0.1, 0.2]) & \text{with probability } 1/2, \\ U([0.1, 1]) & \text{with probability } 1/2. \end{cases}$$



(a) (p9) with $|V_{\alpha>0}|/|V| = 0$ (b) (p10) with $|V_{\alpha>0}|/|V| = 0.2$

Fig. 8: Examples showing that the Σ -model (6) is not globally attractive when g_v is not subhomogeneous.

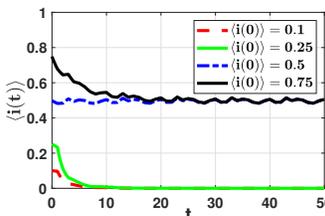


Fig. 9: Example with parameter set (p11) showing that the Σ -model (6) is not globally attractive when $\{\beta_v(t)\}_{t \geq 0}$ is not ergodic.

Note that $\{\beta_v(t)\}_{t \geq 0}$ is not ergodic in this example. Figure 9 plots the dynamics of $\langle i(t) \rangle$ with parameter set (p11) but different initial values $\langle i(0) \rangle$ as well as different realizations of the non-ergodic $\{\beta_v(t)\}_{t \geq 0}$. We observe that the dynamics is not globally attractive because different initial values lead to different trajectories, hinting that ergodicity may be necessary.

V. CONCLUSION

We have proved that preventive and reactive cyber defense dynamics with ergodic time-dependent parameters is globally attractive and the dynamics is further (almost) periodic when the time-dependent parameters are (almost) periodic. These theoretical results supersede the state-of-the-art understanding of at least two models extensively investigated in the literature, and shed a light on the boundary between “when the dynamics is analytically treatable” and “when the dynamics is not analytically treatable”. There are important, but challenging, open problems for future research. First, we numerically showed that ergodicity may be necessary for global attractivity. It is therefore important to rigorously pin down the necessary conditions under which the dynamics is globally attractive, namely the precise boundary between “when the dynamics is analytically treatable” and “when the dynamics is not analytically treatable”. Second, we did not characterize the convergence speed, which is another challenging problem because the globally attractive trajectory is time-dependent, rendering the eigenvalue analysis of Jacobian matrix not applicable here.

REFERENCES

- [1] S. Xu, “Cybersecurity dynamics,” in *Proc. HotSoS’14*, 2014, pp. 14:1–14:2.
- [2] —, “Emergent behavior in cybersecurity,” in *Proc. HotSoS’14*, 2014, pp. 13:1–13:2.
- [3] —, *Cybersecurity Dynamics: A Foundation for the Science of Cybersecurity*. Springer, 2019, pp. 1–31.

- [4] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *IEEE Symp. on Security and Privacy*, 1991, pp. 343–359.
- [5] J. Kephart and S. White, "Measuring and modeling computer virus prevalence," in *IEEE Symp. on Security and Privacy*, 1993, pp. 2–15.
- [6] W. Kermack and A. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. Math. Phys. Eng. Sci.*, vol. 115, pp. 700–721, 1927.
- [7] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *Proc. IEEE SRDS'03*, 2003, pp. 25–34.
- [8] A. Ganesh, L. Massoulie, and D. Towsley, "The effect of network topology on the spread of epidemics," in *Proc. IEEE INFOCOM'05*, 2005, pp. 1455–1466.
- [9] X. Li, T. Parker, and S. Xu, "Towards quantifying the (in)security of networked systems," in *Proc. IEEE AINA'07*, 2007, pp. 420–427.
- [10] X. Li, P. Parker, and S. Xu, "A stochastic model for quantitative security analyses of networked systems," *IEEE Trans. Dependable Sec. Comput.*, vol. 8, no. 1, pp. 28–43, 2011.
- [11] S. Xu, W. Lu, and L. Xu, "Push- and pull-based epidemic spreading in arbitrary networks: Thresholds and deeper insights," *ACM Trans. Autonom. Adapt. Syst.*, vol. 7, no. 3, pp. 32:1–32:26, 2012.
- [12] R. Zheng, W. Lu, and S. Xu, "Preventive and reactive cyber defense dynamics is globally stable," *IEEE Trans. Netw. Sci. Eng.*, vol. 5, no. 2, pp. 156–170, 2018.
- [13] Z. Lin, W. Lu, and S. Xu, "Unified preventive and reactive cyber defense dynamics is still globally convergent," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1098–1111, 2019.
- [14] W. Lu, S. Xu, and X. Yi, "Optimizing active cyber defense dynamics," in *Proc. GameSec'13*, 2013, pp. 206–225.
- [15] S. Xu, W. Lu, and H. Li, "A stochastic model of active cyber defense dynamics," *Internet Math.*, vol. 11, no. 1, pp. 23–61, 2015.
- [16] R. Zheng, W. Lu, and S. Xu, "Active cyber defense dynamics exhibiting rich phenomena," in *Proc. HotSoS'15*, 2015, pp. 2:1–2:12.
- [17] G. Da, M. Xu, and S. Xu, "A new approach to modeling and analyzing security of networked systems," in *Proc. HotSoS'14*, 2014, pp. 6:1–6:12.
- [18] S. Xu, W. Lu, L. Xu, and Z. Zhan, "Adaptive epidemic dynamics in networks: Thresholds and control," *ACM Trans. Autonom. Adapt. Syst.*, vol. 8, no. 4, p. 19, 2014.
- [19] Y. Han, W. Lu, and S. Xu, "Characterizing the power of moving target defense via cyber epidemic dynamics," in *Proc. HotSoS'14*, 2014, pp. 10:1–10:12.
- [20] N. Provos, D. McNamee, P. Mavrommatis, K. Wang, and N. Modadugu, "The ghost in the browser analysis of web-based malware," in *Proc. HotBots'07*, 2007.
- [21] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 1–14, Feb. 2009.
- [22] P. Van Mieghem, "The N-intertwined SIS epidemic network model," *Computing*, vol. 93, pp. 147–169, 2011.
- [23] W. K. Chai and G. Pavlou, "Path-based epidemic spreading in networks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 1, pp. 565–578, Feb. 2017.
- [24] A. Fall, A. Iggidr, G. Sallet, and J.-J. Tewa, "Epidemiological models and Lyapunov functions," *Math. Model Nat. Phenom.*, vol. 2, no. 1, pp. 62 – 83, 2007.
- [25] A. Khanafer, T. Basar, and B. Gharesifard, "Stability properties of infected networks with low curing rates," in *proc. 2014 ACC*, 2014, pp. 3579–3584.
- [26] F. D. Sahneh, C. Scoglio, and P. Van Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1609–1620, Oct. 2013.
- [27] Y.-Q. Zhang, X. Li, and A. V. Vasilakos, "Spectral analysis of epidemic thresholds of temporal networks," *IEEE trans. Cybern.*, no. 99, pp. 1–13, 2017.
- [28] B. A. Prakash, H. Tong, N. Valler, M. Faloutsos, and C. Faloutsos, "Virus propagation on time-varying networks: Theory and immunization algorithms," in *proc. ECML PKDD*, 2010, pp. 99–114.
- [29] M. R. Sanatkar, W. N. White, B. Natarajan, C. M. Scoglio, and K. A. Garrett, "Epidemic threshold of an SIS model in dynamic switching networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 3, pp. 345–355, 2016.
- [30] V. Bokharaie, O. Mason, and F. Wirth, "Spread of epidemics in time-dependent networks," in *Proc. 19th MTNS*, vol. 5, no. 9, 2010.
- [31] P. E. Paré, C. L. Beck, and A. Nedić, "Epidemic processes over time-varying networks," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1322–1334, 2018.
- [32] M. Pendleton, R. Garcia-Lebron, J. Cho, and S. Xu, "A survey on systems security metrics," *ACM Comput. Surv.*, vol. 49, no. 4, pp. 62:1–62:35, 2017.
- [33] M. A. Rami, V. S. Bokharaie, O. Mason, and F. R. Wirth, "Stability criteria for SIS epidemiological models under switching policies," *DCDS-B*, vol. 19, pp. 2865–2887, 2014.
- [34] M. Ogura and V. M. Preciado, "Disease spread over randomly switched large-scale networks," in *proc. 2015 ACC*. IEEE, 2015, pp. 1782–1787.
- [35] L. Arnold, *Random dynamical systems*. Springer, 2013.
- [36] G. D. Birkhoff, "Proof of the ergodic theorem," *PNAS*, vol. 17, no. 12, pp. 656–660, 1931.
- [37] A. A. Borovkov, *Ergodicity and stability of stochastic processes*. J. Wiley, 1998.
- [38] A. Besicovitch, *Almost periodic functions*. Cambridge, 1932.
- [39] M. W. Hirsch, "Systems of differential equations that are competitive or cooperative ii: Convergence almost everywhere," *SIAM J. Math. Anal.*, vol. 16, no. 3, pp. 423–439, 1985.
- [40] X.-Q. Zhao, *Dynamical systems in population biology*. Springer, 2003.
- [41] H. I. Freedman, S. Ruan, and M. Tang, "Uniform persistence and flows near a closed positively invariant set," *J. Dynam. Differ. Equat.*, vol. 6, no. 4, pp. 583–600, 1994.

- [42] G. Butler, H. I. Freedman, and P. Waltman, “Uniformly persistent systems,” *Proc. Am. Math. Soc.*, pp. 425–430, 1986.
- [43] L. Moreau, “Stability of continuous-time distributed consensus algorithms,” in *proc. 43rd IEEE CDC*, vol. 4, 2004, pp. 3998–4003.
- [44] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. Siam, 1994, vol. 9.
- [45] Y. Han, W. Lu, and T. Chen, “Achieving cluster consensus in continuous-time networks of multi-agents with inter-cluster non-identical inputs,” *IEEE Trans. Automat. Contr.*, vol. 60, no. 3, pp. 793–798, 2015.
- [46] —, “Cluster consensus in discrete-time networks of multiagents with inter-cluster nonidentical inputs,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 4, pp. 566–578, 2013.
- [47] H. L. Smith, *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems*. American Mathematical Society, 2008, no. 41.
- [48] W. Lu and T. Chen, “Almost periodic dynamics of a class of delayed neural networks with discontinuous activations,” *Neural Comput.*, vol. 20, no. 4, pp. 1065–1090, 2008.
- [49] R. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos (2dn Edition)*. CRC Press, 1999.

APPENDIX A
PROOF OF LEMMA 1

Proof. For proving (i), we observe that the boundedness of $B(t)$ means that there exists $M_1 > 0$ such that $\sum_{k=1}^n b_{jk}(t) \geq -M_1$ holds for $\forall j, \forall t$. Let $z_{i_0(t)}(t) = \min_i z_i(t)$. It follows

$$\begin{aligned} \frac{dz_{i_0(t)}}{dt} &= b_{i_0(t)i_0(t)}(t)z_{i_0(t)}(t) + \sum_{j \neq i_0(t)} b_{i_0(t)j}(t)z_j(t) \\ &\geq b_{i_0(t)i_0(t)}(t)z_{i_0(t)}(t) + \sum_{j \neq i_0(t)} b_{i_0(t)j}(t)z_{i_0(t)}(t) \\ &\geq -M_1 z_{i_0(t)}(t); \end{aligned} \quad (10)$$

this implies $z_{i_0(t)}(t) \geq 0$ when $z(0) \geq \mathbf{0}$. This means that the solution matrix $U(t, s)$ is nonnegative for any $t \geq s \geq 0$. It follows from the ergodicity of $\{B(t)\}_{t \in \mathbb{R}}$ that the convergence

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_a^{a+t} B(s) ds = M(B) \quad (11)$$

holds uniformly for any $a \in \mathbb{R}$.

For proving (ii), we observe that if $\mathcal{G}(M(B))$ is strongly connected, then there exists $\delta > 0$ such that δ -graph of $M(B)$ is strongly connected. It then follows from Eq. (11) that there exists $T_1 > 0$, such that for any interval $[kT_1, (k+1)T_1)$ for any $k \in \mathbb{N}$, the $\frac{\delta}{2}$ -graph of $\int_{kT_1}^{(k+1)T_1} B(s) ds$ is strongly connected. According to the proof of Lemma 2(2) in [45], there exists $\eta_1 > 0$ such that the η_1 -matrix of $U((k+1)T_1, kT_1)$ is irreducible. On the other hand, Inequality (10) implies that

$$U_{jj}(t, s) \geq e^{-M_1(t-s)}, \forall j. \quad (12)$$

Following the proof of Lemma 1 in [46], we know that the product of $(n-1)$ -many n -dimensional nonnegative matrices, which are irreducible and have positive diagonal elements, gives a positive matrix. This means that there exists $\eta_2 > 0$ such that

$$U(t_{k+n}, t_{k+1}) = U(t_{k+n}, t_{k+n-1}) \cdots U(t_{k+2}, t_{k+1}) > \eta_2 \quad (13)$$

for $\forall k \in \mathbb{N}$. Let $T = nT_1$ and $t_k = kT$, for any $k \in \mathbb{N}$. Then by Inequalities (12) and (13), the following holds for any $\Delta > T$ and any $s \in [t_N, t_{N+1})$, $\forall N \in \mathbb{N}$:

$$\begin{aligned} U(s + \Delta, s) &= U(t_{N+1}, s)U(t_{N+n}, t_{N+1})U(s + \Delta, t_{N+n}) \\ &\geq \eta_2 e^{-M_1(t_{N+1}-s)} e^{-M_1(s+\Delta-t_{N+n})} \geq \eta_2 e^{-M_1 \Delta}. \end{aligned}$$

By letting $\epsilon = \eta_2 e^{-M_1 \Delta}$, we complete the proof. \square

APPENDIX B
PROOF OF THEOREM 3

Proof. It follows from Property 3 and Theorem 2 that $f(i, y)$ is strongly subhomogeneous, meaning that for any $\alpha > 0$ and $i \in [0, 1]^n$, we have

$$\begin{aligned} f(i, y) &\leq \frac{f(\alpha i, y)}{\alpha} = \frac{f(\mathbf{0}, y) + \alpha D_i f(\mathbf{0}, y) i + o(\|\alpha i\|)}{\alpha} \\ &= D_i f(\mathbf{0}, y) i + \frac{o(\|\alpha i\|)}{\alpha}. \end{aligned}$$

When $\alpha \rightarrow 0$, we have $f(i, y) \leq D_i f(\mathbf{0}, y) i$. This means that for $i \in \mathbb{R}_{\geq 0}^n$, system $dz(t)/dt = D_i(\mathbf{0}, y(t))z(t)$ is a

comparison model to model (4). If $i(0) = z(0) \in [0, 1]^n$, $i_v(t) \leq z_v(t)$ holds for $\forall t > 0, \forall v$.

When $\mu(D_i f(\mathbf{0}, y)) < 0$, it follows that $\forall v$, $\lim_{t \rightarrow \infty} z_v(t) = 0$ holds for any initial value $z(0) \in [0, 1]^n$. The global convergence to the equilibrium $\mathbf{0}$ follows from the fact $i_v(t) \leq z_v(t)$. \square

APPENDIX C
PROOF OF LEMMA 2

Proof. Consider the linear variational equation of model (4) at the equilibrium $\mathbf{0}$,

$$\frac{dz}{dt} = D_i f(\mathbf{0}, y(t))z, \quad (14)$$

where $D_i \psi(t, \mathbf{0}, y)$ is the fundamental solution matrix. Since $\{y(t)\}_{t \geq 0}$ is ergodic and the Jacobian matrix $D_i f(\mathbf{0}, y)$ is bounded, Oseledets multiplicative ergodic theorem of random dynamical systems (Theorems 3.4.1 and 3.4.11 in [35]) says that there exists an invariant set $\tilde{Y} \in \mathcal{F}$ of full measure on which there is an Oseledets splitting $\mathbb{R}^n = E_1(y) \oplus \cdots \oplus E_r(y)$ associated with $\rho_1 > \cdots > \rho_r$ such that for $k = 1, \dots, r$,

$$D_i \psi(t, \mathbf{0}, y) E_k(y) = E_k(\theta(t, y))$$

and

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|D_i \psi(t, \mathbf{0}, y) \xi\| = \rho_k, \text{ for } \xi \in E_k(y) \setminus \{\mathbf{0}\}, \quad (15)$$

where convergence is uniform in $E_k(y) \cap S_1^n$ with $S_1^n = \{x \in \mathbb{R}^n : \|x\| = 1\}$. On the other hand, it follows from Properties 1-2 and Lemma 1 that $D_i \psi(t, \mathbf{0}, y)$ is nonnegative for all $t \in \mathbb{R}$, $y \in Y$ and $\mathbb{R}_{\geq 0}^n$ is invariant with respect to model (14). Hence, $E_1(y) \subseteq \mathbb{R}_{\geq 0}^n$.

Consider the dual system of model (14) with $s < t$ as follows

$$\frac{dx(s)}{ds} = -x(s) D_i f(\mathbf{0}, y(s)), \quad (16)$$

to which the fundamental solution matrix is $U_1(s, t)$. Then, we have $U_1(s, t) = U(t, s)$. It can be seen that model (16) has Lyapunov exponents $-\rho_j, j = 1, \dots, r$. Denote the associated Oseledets splitting subspaces by $F_j(y), j = 1, \dots, r$. Since $\mathbb{R}_{\leq 0}^n$ is invariant for model (16), $F_1(y) \in \mathbb{R}_{\leq 0}^n$ for all x and y .

Consider $\varphi_1^\top(t) U(t, s) \phi_j(s)$ with $\phi_j(s) \in E_j(\theta(s, y)) \cap S_1^n$ and $\varphi_1^\top(t) \in F_1(\theta(t, y)) \cap S_1^n$ for any $j \neq 1$. As $t-s \rightarrow +\infty$, it can be respectively approximated as follows:

$$\begin{aligned} \varphi_1^\top(t) U(t, s) \phi_j(s) &\sim \exp(\rho_j(t-s)) \varphi_1^\top(t) \phi_j(t), \\ \varphi_1^\top(t) U(t, s) \phi_j(s) &\sim \exp(-\rho_1(s-t)) \varphi_1^\top(s) \phi_j(s), \end{aligned}$$

which implies

$$\varphi_1^\top(t) \phi_j(t) \sim \exp((\rho_1 - \rho_j)(t-s)) \varphi_1^\top(s) \phi_j(s).$$

Since $\varphi_1(t)$ and $\phi_j(t)$ are bounded, we have $\varphi_1^\top(t) \phi_j(t) = 0$ for all t and $j \neq 1$. On the other hand, from Property 2 and $\mathcal{G}(M(\Gamma))$ being strongly connected, we know $\mathcal{G}(M(D_i f(\mathbf{0}, y)))$ is strongly connected. Then, it follows from Lemma 1 that $U(t, s)$ is non-negative and there exist $\Delta > 0$ and $\epsilon_1 > 0$ such that every element of $U(t + \Delta, t)$ is greater

than ϵ_1 . This means that $\varphi_1(t) \in \mathbb{R}_{\leq 0}^n \cap S_1^n$ when $t \geq \Delta$. Then, from the compactness of $[0, 1]^n$ and Y , we have that the angles $\angle(E_j(y), \mathbb{R}_{>0}^n) > \eta$ for $j \neq 1$ and some $\eta > 0$. Then, we have that there exists $M > 0$ such that for any $\xi \in \mathbb{R}_{>0}^n$,

$$\|D_i \psi(t, \mathbf{0}, y) \xi\| \geq M \|\xi\| \exp(rt)$$

holds for some $r > 0$. Since Property 1 says that $f(i, y)$ has continuous first and second derivatives with respect to i and model (14) is the first order linear approximation to model (4) around the equilibrium $\mathbf{0}$, there exists $\epsilon > 0$ and $T' > 0$ such that for any $i_0 \in \mathbb{R}_{>0}^n$, we have

$$\|\psi(t, i_0, y)\| > \epsilon \quad (17)$$

when $t > T'$. For an index set $C \subseteq \{1, \dots, n\}$, let

$$Y_\epsilon^C = \{x \in \mathbb{R}_{\geq 0}^n : x_j = 0, \forall j \in C \text{ and } x_{j'} > \epsilon, \forall j' \notin C\}.$$

Suppose $\psi(t, i_0, y)$ goes near Y_ϵ^C and let $z(t) = [i(t)]_C \in \mathbb{R}^{|C|}$, it follows from the mean value theorem that

$$\frac{dz(t)}{dt} = A(t)z(t) + I(t), \quad (18)$$

where $A(t) = [D_i f(\hat{i}, y)]_{C,C}$ and $I(t) = [f(\hat{i}, y)]_C$ with $\hat{i} \in Y_\epsilon^C$ and $\hat{i} = \tilde{c} + (1 - c)i$, $c \in (0, 1)$. From the definition of Y_ϵ^C and the strong connectivity of $\mathcal{G}(M(\Gamma))$, we know there exists $T_1 > 0$ such that $\max_j \int_t^{t+T_1} I_j(t) dt > \epsilon'$ for some $\epsilon' > 0$ and any $t > 0$. Let $U_2(t, s)$ denote the fundamental solution matrix of system $du(t)/dt = A(t)u(t)$. Lemma 1 says that there exists $T_2 > 0$ such that for each $\Delta > T_2$, one can find $\epsilon''(\Delta) > 0$ such that $U_2(t + \Delta, t) \geq \epsilon''(\Delta)$. This means that the solution to model (18) with non-negative initial value, denoted by $z(t)$, satisfies the following: For $t > T_1 + T_2 + 1$,

$$\begin{aligned} z(t) &\geq \int_0^t U_2(t, s) I(s) ds \geq \int_{t-T_1-T_2}^{t-T_2} \epsilon''(t-s) I(s) ds \\ &\geq \min_{\Delta \in [T_2, T_1+T_2]} \epsilon''(\Delta) \int_{t-T_1-T_2}^{t-T_2} I(s) ds \geq \epsilon' \min_{\Delta \in [T_2, T_1+T_2]} \epsilon''(\Delta). \end{aligned}$$

This implies that for each $C \subset \{1, \dots, n\}$ and a sufficiently small $\epsilon > 0$, there exists $v_{C,\epsilon} > 0$ such that for each $i_{C,0} \in Y_\epsilon^C$, there exists $T > 0$ so that $\|\psi(t, i_{C,0}, y)\|_C > v_{C,\epsilon}$ for each $t > T$.

Now we can complete the proof of persistence by induction. First, inequality (17) implies that the trajectory $\psi(t, i_0, y)$ essentially goes out of the ball $B_\epsilon^n = \{x \in \mathbb{R}_{\geq 0}^n : \|x\| \leq \epsilon\}$. So, there exists at least one index, say $j(t)$, such that $[\psi(t, i_0, y)]_{j(t)} > \epsilon_1 = \epsilon/n$ for any sufficiently large t . If $\psi(t, i_0, y)$ goes near $Y_{\epsilon_1}^{C_j}$ where $C_j = \{1, \dots, n\} - \{j(t)\}$, the preceding analysis indicates that $[\psi(t, i_0, y)]_{C_j}$ essentially goes out of the ball $B_{\epsilon_1}^{C_j} = \{x : \|x_{C_j}\| \leq \epsilon_1\}$.

Then, by induction, we can prove that $\psi(t, i_0, y)$ essentially goes out of the union of the following balls

$$\bigcup_{C \subset \{1, \dots, n\}} B_{\epsilon_C}^C,$$

meaning that $\psi(t, i_0, y)$ is persistent for any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$. Now it can be concluded that $\psi(t, i_0, y)$ is persistent for any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$ and almost all $y \in Y$. \square

APPENDIX D PROOF OF LEMMA 3

Proof. First, we prove part (i). Let Y denote the parameter space. Note that when $y \in Y$ is ergodic and Y is compact, the conditions on Y in Theorem 1 are satisfied. Therefore, the global attractivity result can be derived by employing Theorem 1 when the dynamics is uniformly persistent and satisfies Conditions (I) and (II) in Theorem 1. Since the uniform persistence of the dynamics has been proven in Lemma 2, we only need to prove that the dynamics satisfies Conditions (I) and (II) in Theorem 1.

From Property 3 and Theorem 2, we know $f(i, y)$ is strongly subhomogeneous. For any $\eta \in (0, 1)$, let $u_1(t) = \eta \psi(t, i_0, y)$ and $u_2(t) = \psi(t, \eta i_0, y)$. Then, $u_1(t)$ satisfies

$$\begin{aligned} \frac{du_1(t)}{dt} &= \eta \frac{d\psi(t, i_0, y)}{dt} = \eta f(\psi(t, i_0, y), y) \\ &\leq f(\eta \psi(t, i_0, y), y) = f(u_1(t), y) \end{aligned}$$

and $u_2(t)$ satisfies $du_2(t)/dt = f(u_2(t), y)$. From the comparison theory of differential equations and $u_1(0) = u_2(0) = \eta i_0$, we have that $u_1(t) \leq u_2(t)$, i.e. $\psi(t, \eta i_0, y) \geq \eta \psi(t, i_0, y)$ for $\forall t, i_0, y$. Hence, we know that for any $t \in \mathbb{R}$ and $y \in Y$, $\psi(t, \cdot, y)$ is subhomogeneous on $[0, 1]^n$. Recall that we have proved that $\psi(t, i_0, y)$ is uniformly persistent, namely that there exists T_2 such that $\psi(t, i_0, y) \gg \mathbf{0}$ for $t \geq T_2$ and any i_0, y . By the strong subhomogeneity of $f(\cdot, y)$, it follows that $u_1(t_0) \ll u_2(t_0)$ for some $t_0 \geq T_2$, implying that for any $y \in Y$, $\psi(t_0, \cdot, y)$ is strongly subhomogeneous on $[0, 1]^n$.

Note that Theorem 4.1.1 in [47] proved the monotonicity for cooperative and irreducible dynamical systems. Here we extend this result to our model (4). For each $(i, y) \in [0, 1]^n \times Y$, let $Z(t) = \partial \psi(t, i, y) / \partial i$. Then, we have

$$\frac{dZ(t)}{dt} = B(t)Z(t), \quad \text{where } B(t) = D_i f(i, y). \quad (19)$$

It follows from Lemma 1 that every element of $Z(t)$ is nonnegative. Then, the monotonicity of $\psi(t, \cdot, y)$ follows from

$$\begin{aligned} &\psi(t, \hat{i}_0, y) - \psi(t, i_0, y) \\ &= \int_0^1 \frac{\partial \psi}{\partial i}(t, i_0 + r(\hat{i}_0 - i_0), y)(\hat{i}_0 - i_0) dr. \end{aligned} \quad (20)$$

For any $i_0, \hat{i}_0 \gg \mathbf{0}$, we can always find $\eta \in (0, 1)$ such that $\eta i_0 \leq \hat{i}_0 \leq \eta^{-1} i_0$. Then, by the subhomogeneity and monotonicity of $\psi(t, \cdot, y)$ on $[0, 1]^n$, we know that for any $t \in \mathbb{R}$,

$$\eta \psi(t, i_0, y) \leq \psi(t, \hat{i}_0, y) \leq \eta^{-1} \psi(t, i_0, y).$$

From the strongly subhomogeneity of $\psi(t_0, \cdot, y)$, it follows

$$\eta \psi(t_0, i_0, y) \ll \psi(t_0, \hat{i}_0, y) \ll \eta^{-1} \psi(t_0, i_0, y).$$

This means Conditions (I) and (II) in Theorem 1 are satisfied. This completes the proof of part (i).

Now we prove part (ii). From $\alpha_v(t) = 0$, $h_v(i(t), \beta_v(t)) \leq 1$, $g_v(i(t), \alpha_v(t), \Gamma(t)) \geq 0$ for $\forall v \in V$ and $t \geq 0$, we know $\frac{di_v(t)}{dt} \geq -i_v(t)$ always holds, implying $i_v(t) \geq \exp(-t)i_v(0)$. It follows that when $i_v(0) > 0$, $i_v(t) > 0$ holds for any $t \geq 0$. From the strong connectivity of $\mathcal{G}(M(\Gamma))$ and Property 2, we

know there exists $T_1 > 0$ such that $i_v(T_1) > 0$ for any $v \in V$. Therefore, for any $i_0, \hat{i}_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$, we can find $\eta \in (0, 1)$ such that

$$\eta\psi(T_1, i_0, y) \leq \psi(T_1, \hat{i}_0, y) \leq \eta^{-1}\psi(T_1, i_0, y). \quad (21)$$

By combing with the subhomogeneous and monotonicity of $\psi(t, \cdot, y)$ on $[0, 1]^n$, we have that for $t \geq T_1$,

$$\eta\psi(t, i_0, y) \leq \psi(t, \hat{i}_0, y) \leq \eta^{-1}\psi(t, i_0, y). \quad (22)$$

From the strong subhomogeneity of solution $\psi(t_0, \cdot, y)$ proven in the preceding part (i), we have that for any $n \geq 1, n \in \mathbb{N}$,

$$\begin{aligned} \eta\psi(nT_0 + T_1, i_0, y) &\ll \psi(nT_0 + T_1, \hat{i}_0, y) \\ &\ll \eta^{-1}\psi(nT_0 + T_1, i_0, y). \end{aligned} \quad (23)$$

We recall the metric $\rho(\cdot, \cdot)$ given in [40]:

$$\begin{aligned} \rho(i_0, \hat{i}_0) &= \inf\{\ln \eta \mid \eta \geq 1, \eta^{-1}i_0 \leq \hat{i}_0 \leq \eta i_0\}, \\ &\quad \forall i_0, \hat{i}_0 \in (0, 1]^n. \end{aligned}$$

Let $X = [0, 1]^n$. For any $y \in Y$, we define $\hat{\rho} : X \rightarrow \mathbb{R}^+$ as

$$\hat{\rho}(i_0, \hat{i}_0) = \rho(\psi(T_1, i_0, y), \psi(T_1, \hat{i}_0, y))$$

for $\forall i_0, \hat{i}_0 \in X \setminus \{\mathbf{0}\}$, $\hat{\rho}(i_0, \mathbf{0}) = +\infty$ and $\hat{\rho}(\mathbf{0}, \mathbf{0}) = 0$. It can be seen that $(X, \hat{\rho})$ is a metric space and $\hat{\rho}(\cdot, \cdot)$ is continuous with respect to the product topology induced by the norm $\|\cdot\|$. Moreover, inequalities (21) and (22) imply that for any $i_0, \hat{i}_0 \in X \setminus \{\mathbf{0}\}$ and $t \geq T_1$,

$$\hat{\rho}(\psi(t, i_0, y), \psi(t, \hat{i}_0, y)) \leq \hat{\rho}(i_0, \hat{i}_0), \quad (24)$$

while inequalities (21) and (23) imply that for any $i_0, \hat{i}_0 \in X \setminus \{\mathbf{0}\}$ with $i_0 \neq \hat{i}_0$ and $n \geq 1, n \in \mathbb{N}$,

$$\hat{\rho}(\psi(nT_0, +T_1, i_0, y), \psi(nT_0, +T_1, \hat{i}_0, y)) < \hat{\rho}(i_0, \hat{i}_0).$$

Since X and Y are compact, we have that any trajectory $(\psi(t, i_0, y), \theta(t, y))$ has its ω -limit set $K_0 \subseteq X \times Y$.

If $K_0|_X \neq \{\mathbf{0}\}$, similar to the proof of Theorem 2.3.5 in [40] (Theorem 1 in the present paper), we can prove that for any $y_0 \in K_0|_Y$, the cardinality of set $p^{-1}(y_0) \cap K_0$ is one, i.e., there is only one point $i^* \in K_0|_X$ satisfies $(i^*, y_0) \in K_0$. For any trajectory with any nonzero initial value, its ω -limit set equals K_0 . This implies that for any $y \in Y$, model (4) is globally attractive.

If there exists $y_0 \in Y$ such that $\psi(t, i_{y_0}^*, y_0) \in (0, 1]^n$, then $K_0|_X \subseteq (0, 1]^n$. Notice that for any trajectory with any nonzero initial value, its ω -limit set equals K_0 , implying that $\psi(t, i_{y_0}^*, y) \in (0, 1]^n$ holds for almost every $y \in Y$.

If $K_0|_X = \{\mathbf{0}\}$, then by inequality (24) and the assumption that $\hat{\rho}(\mathbf{0}, i_0) = +\infty$, we have that for any trajectory with any nonzero initial value, its omega limit set equals K_0 . This implies that the equilibrium $\mathbf{0}$ is globally attractive.

Finally, part (iii) follows from Theorem 3 immediately. \square

APPENDIX E PROOF OF LEMMA 4

Proof. From $h_v(i(t), \beta_v(t)) \leq 1$ and $g_v(i(t), \alpha_v(t), \Gamma(t)) \geq \alpha_v(t)$, we know that for $t \geq 0$,

$$\frac{di_v(t)}{dt} \geq -i_v(t) + \alpha_v(t)(1 - i_v(t)) \geq -2i_v(t) + \alpha_v(t).$$

It follows from $M(\alpha_v) > 0$ that there exist $\delta_1 > 0$ and $T_0 > 0$ such that

$$\int_{kT_0}^{(k+1)T_0} \alpha_{v_0}(s) ds \geq \delta_1$$

holds for any $k \in \mathbb{N}$. Then for $t \in [NT_0, (N+1)T_0)$ and $N \in \mathbb{N}$, we have

$$\begin{aligned} i_{v_0}(t) &\geq e^{-2t}i_{v_0}(0) + \int_0^t e^{2(s-t)}\alpha_{v_0}(s) ds \\ &\geq \sum_{k=0}^N e^{-2t} \int_{2kT_0}^{(k+1)T_0} e^{kT_0}\alpha_{v_0}(s) ds \\ &\geq \frac{e^{-2T_0} - e^{-2(N+1)T_0}}{e^{2T_0} - 1} \delta_1. \end{aligned}$$

Therefore, the trajectory of node v_0 is persistent. Since $\mathcal{G}(M(\Gamma))$ is strongly connected, we know that there exist paths from v_0 to other nodes in $\mathcal{G}(M(\Gamma))$. Then, we have that there exist node $u_0 \in V$, $\delta > 0$ and $T > 0$ such that $\int_{t_k}^{t_{k+1}} \gamma_{u_0 v_0} ds \geq \delta$, where $t_k = kT, k \in \mathbb{N}$. It then follows from Property 2 that there exists $\delta' > 0$ such that for any $k \in \mathbb{N}$,

$$\int_{t_k}^{t_{k+1}} \frac{\partial g_{u_0}}{\partial i_{v_0}}(\cdot, \alpha_{v_0}(s), \Gamma(s)) ds \geq \delta'. \quad (25)$$

Then, from $h_{u_0} \leq 1$, $\partial g_v / \partial i_u \geq 0$ and the integral mean value theorem, we have

$$\begin{aligned} \frac{di_{u_0}(t)}{dt} &\geq -h_{u_0}i_{u_0} + g_{u_0}(1 - i_{u_0}) \\ &= -(h_{u_0} + g_{u_0})i_{u_0} + \left[\alpha_{u_0} + \frac{\partial g_{u_0}}{\partial i_{v_0}}(\hat{i}, \alpha_{v_0}, \Gamma)i_{v_0} \right. \\ &\quad \left. + \sum_{v \neq v_0} \frac{\partial g_{u_0}}{\partial i_v}(\hat{i}, \alpha_{v_0}, \Gamma)i_v \right] \\ &\geq -2i_{u_0} + \alpha_{u_0} + \frac{\partial g_{u_0}}{\partial i_{v_0}}(\hat{i}, \alpha_{v_0}, \Gamma)i_{v_0}. \end{aligned}$$

By combing with Inequality (25), we can see that there exists T_{u_0} such that $i_{u_0}(t) \geq \delta_{u_0}$ for $t \geq T_{u_0}$. By induction, we can prove that for any $v \in \{1, \dots, n\}$, there exists $\delta'' > 0$ and $T' > 0$ such that $i_v(t) \geq \delta''$. In other words, all trajectories are uniformly persistent. On the other hand, it follows from the proof of Lemma 3 that $\psi(t, \cdot, y)$ satisfies Conditions (I) and (II) in Theorem 1, meaning that model (4) is globally attractive. \square

APPENDIX F PROOF OF LEMMA 5

Proof. First, we prove part (i) of Lemma 5. The global attractivity for the dynamics corresponding to strongly connected component SCC_1 can be derived from Lemmas 3 and 4 directly. Let $\phi(t, i_{SCC_1}^*, y)$ denote the globally attractive

trajectory corresponding to SCC_1 . For any given ergodic function $z(t) \in [0, 1]^{|SCC_1|}$, the system

$$\frac{di_v(t)}{dt} = f_v([z(t), [i(t)]_{SCC_2}], y), \quad v \in SCC_2 \quad (26)$$

has a globally attractive trajectory in $[0, 1]^{|SCC_2| \setminus \{0\}}$ by treating $(y(t), z(t))$ as enlarged parameters.

Let $\{\epsilon_k\}_{k \in \mathbb{N}}$ be a positive and monotone decreasing sequence that converges to 0. Then, by the global attractivity of SCC_1 , we have that for any $i_{2,0} \in [0, 1]^{|SCC_2|}$ and any ϵ_k , there exists $T_k > 0$ such that for $t > T_k$ and any $i_{1,0} \in [0, 1]^{|SCC_1|}$,

$$\begin{aligned} & \max\{\mathbf{0}_{SCC_1}, \phi(t, i_{SCC_1}^*, y) - \epsilon_k \mathbf{1}_{SCC_1}\} \\ & \leq [\psi(t, [i_{1,0}, i_{2,0}], y)]_{SCC_1} \leq \phi(t, i_{SCC_1}^*, y) + \epsilon_k \mathbf{1}_{SCC_1}. \end{aligned}$$

For brevity, denote

$$\begin{aligned} i(t) &= \psi(t, [i_{1,0}, i_{2,0}], y), \\ i_{k,1}^l(t) &= \max\{\mathbf{0}_{SCC_1}, \phi(t, i_{SCC_1}^*, y) - \epsilon_k \mathbf{1}_{SCC_1}\}, \\ i_{k,1}^u(t) &= \phi(t, i_{SCC_1}^*, y) + \epsilon_k \mathbf{1}_{SCC_1}. \end{aligned}$$

Let $i_{k,2}^q(t)$ denote the globally attractive trajectory of system (26) with $z(t) = i_{k,1}^q(t)$, $q = l, u$, respectively. According to Theorem 1.7 in [39], a cooperative function is monotone, meaning that $f_v(i, y)$ is monotone in i , which implies

$$\begin{aligned} f_v([i_{k,1}^l(t), [i(t)]_{SCC_2}], y) &\leq f_v(i(t), y) \\ &\leq f_v([i_{k,1}^u(t), [i(t)]_{SCC_2}], y). \end{aligned}$$

Then it can be concluded that $i_{k,2}^l(t) \leq [i(t)]_{SCC_2} \leq i_{k,2}^u(t)$. From $\lim_{k \rightarrow \infty} \|i_{k,1}^u(t) - i_{k,1}^l(t)\| = 0$ and f is continuously differentiable, we can see that $\lim_{k \rightarrow \infty} \|i_{k,2}^u(t) - i_{k,2}^l(t)\| = 0$. Therefore, there exists i^* such that

$$\lim_{k \rightarrow \infty} i_{k,2}^l(t) = \lim_{k \rightarrow \infty} i_{k,2}^u(t) = [\psi(t, i^*, y)]_{SCC_2}$$

for any $i_{1,0} \in [0, 1]^{|SCC_1|}$ and $[\psi(t, i^*, y)]_{SCC_2} = \phi(t, i_{SCC_1}^*, y)$. That is, $\psi(t, i^*, y)$ is the globally attractive trajectory we are seeking for $SCC_1 \cup SCC_2$.

Let $V_{\alpha=0} = \{v : \alpha_v(t) = 0, \forall t \in \mathbb{R}\}$ denote the set of nodes that are not subject to pull-based attacks, meaning that $M(\alpha_v) = 0$. Therefore, V can be partitioned into two sets such that $V = V_{\alpha=0} \cup V_{\alpha>0}$. This means that $\mathcal{V}_{SCC_2} \cap V_{\alpha>0} = \emptyset$ is equivalent to $\mathcal{V}_{SCC_2} \subseteq V_{\alpha=0}$, namely that $\alpha_v(t) = 0, \forall v \in \mathcal{V}_{SCC_2}$. When $\phi(t, i_{SCC_1}^*, y) = \mathbf{0}_{|SCC_1|}$, $\mathcal{V}_{SCC_2} \cap V_{\alpha>0} = \emptyset$ and $\mu(Df_{22}) < 0$, the global attractivity trajectory corresponding to SCC_2 , i.e., $[\psi(t, i^*, y)]_{SCC_2}$ satisfies system (26) with $z(t) = \mathbf{0}_{|SCC_1|}$. Then it follows from part (iii) of Lemma 3 that SCC_2 globally converges to $\mathbf{0}_{|SCC_2|}$. This completed the proof of proves part (i) of Lemma 5.

Now we prove part (ii) of Lemma 5. If $\mathcal{V}_{SCC_2} \cap V_{\alpha>0} \neq \emptyset$ or $\phi(t, i_{SCC_1}^*, y) \in (0, 1]^{|SCC_1|}$, it follows from Lemma 4 that $[\psi(t, i^*, y)]_{SCC_2} \in (0, 1]^{|SCC_2|}$ is globally attractive in $[0, 1]^{|SCC_2|}$. If $\phi(t, i_{SCC_1}^*, y) = \mathbf{0}_{|SCC_1|}$, $\mathcal{V}_{SCC_2} \cap V_{\alpha>0} = \emptyset$ and $\mu(Df_{22}) \geq 0$, by substituting $z(t) = \mathbf{0}_{SCC_1}$ into system (26) and employing Lemma 3, we have that $[\psi(t, i^*, y)]_{SCC_2} \in [0, 1]^{|SCC_2|}$ is globally attractive in $[0, 1]^{|SCC_2| \setminus \{0\}}$. If $\phi(t, i_{SCC_1}^*, y) \notin \{\mathbf{0}_{|SCC_1|}\} \cup (0, 1]^{|SCC_1|}$

and $\mathcal{V}_{SCC_2} \cap V_{\alpha>0} = \emptyset$, similar to the proof of Lemmas 3 and 4, we can prove that there exists a globally attractive trajectory $[\psi(t, i^*, y)]_{SCC_2} \in [0, 1]^{|SCC_2|}$ for SCC_2 . This completes the proof of part (ii) of Lemma 5. \square

APPENDIX G PROOF OF THEOREM 4

Proof. We can always divide the graph $\mathcal{G}(M(\Gamma))$ into K strongly connected components, denoted by SCC_1, \dots, SCC_K . Let V_{SCC_k} denote the set of nodes in SCC_k and $|SCC_k|$ the number of the nodes in SCC_k , $k = 1, \dots, K$. Let R_k be the indices of SCC s that have links pointing to SCC_k . This means that if $R_k \neq \emptyset$ and $j \in R_k$, there exists a node in SCC_j that has a path to a node in SCC_k . From Properties 2 and $\partial f_v(\mathbf{0}, y)/\partial i_v = \partial g_v(\mathbf{0}, \alpha_v, \Gamma)/\partial i_v$, we know that the Jacobian matrix $D_i f(\mathbf{0}, y)$ has the Perron-Frobenius form (9). Following the argument in the proof of Lemma 5, we can use induction to prove that for each SCC_k ,

- $\mathbf{0}_{|SCC_k|}$ is globally attractive in $[0, 1]^n$ if either of the following condition holds:
 - $\mathcal{V}_{SCC_k} \cap V_{\alpha>0} = \emptyset$, $\mu(Df_{kk}) < 0$ and $R_k = \emptyset$;
 - $\mathcal{V}_{SCC_k} \cap V_{\alpha>0} = \emptyset$, $\mu(Df_{kk}) < 0$, $R_k \neq \emptyset$ and $\mathbf{0}_{|SCC_r|}$ is globally attractive for every $r \in R_k$.
- SCC_k has a positive trajectory that is globally attractive in $[0, 1]^{|SCC_k| \setminus \{0\}}$ if one of the following condition holds:
 - $\mathcal{V}_{SCC_k} \cap V_{\alpha>0} = \emptyset$, $\mu(Df_{kk}) > 0$ and $R_k = \emptyset$;
 - $R_k \neq \emptyset$ and there exists $j \in R_k$ such that SCC_j has a positive and globally attractive trajectory.
- SCC_k has a positive trajectory that is globally attractive in $[0, 1]^{|SCC_k|}$ if $\mathcal{V}_{SCC_k} \cap V_{\alpha>0} \neq \emptyset$.
- SCC_k has a trajectory that is globally attractive in $[0, 1]^{|SCC_k| \setminus \{0\}}$ if $\mathcal{V}_{SCC_k} \cap V_{\alpha>0} = \emptyset$ and $\mu(Df_{kk}) \geq 0$.

Therefore, it can be concluded that for $V = \cup_{k=1}^K \mathcal{V}_{SCC_k}$, there exists a globally attractive trajectory $\psi(t, i^*, y) \in [0, 1]^n \setminus \{0\}$. Moreover, following the argument in the proof of Lemma 4, we know that if node v is subject to pull-based attacks, the trajectory of node v is uniformly persistent, implying $[\psi(t, i^*, y)]_v \neq 0$. \square

APPENDIX H PROOF OF THEOREM 5

Proof. For the almost periodic function $y = \{y(t)\}_{t \in \mathbb{R}}$, let Y_0 be the closure of $\{\theta(s, y), s \in \mathbb{R}\}$ and \mathcal{F}_0 be the Borel σ -algebra of Y_0 . The normalized Haar measure of Y_0 , denoted by \mathbb{P}_0 , is the unique θ -invariant and ergodic probability measure [35]. By Definition 1, $\{y(t)\}_{t \in \mathbb{R}}$ is ergodic with respect to the probability space $(Y_0, \mathcal{F}_0, \mathbb{P}_0)$. Then it follows from Theorem 4 that for almost every $y_0 \in Y_0$, there is a globally attractive trajectory in $[0, 1]^n \setminus \{0\}$. Suppose for $y_1 = \theta(t_1, y) \in Y_0$, there is a globally attractive trajectory $\psi(t, i_1^*, y_1)$, namely that $\lim_{t \rightarrow \infty} \|\psi(t, i_0, y_1) - \psi(t, i_1^*, y_1)\| = 0$ for $\forall i_0 \in [0, 1]^n \setminus \{0\}$. Let $i^* = \psi(-t_1, i_1^*, y_1)$. Then,

$$\psi(t_1, i^*, y) = \psi(t_1, \psi(-t_1, i_1^*, y_1), y) = \psi(0, i_1^*, y_1) = i_1^*.$$

For any $i_0 \in [0, 1]^n \setminus \{\mathbf{0}\}$ and $t > t_1$, we have

$$\begin{aligned} & \|\psi(t, i_0, y) - \psi(t, i^*, y)\| \\ &= \|\psi(t - t_1, \psi(t_1, i_0, y), y_1) - \psi(t - t_1, i_1^*, y_1)\|, \end{aligned}$$

implying that $\psi(t, i^*, y)$ is globally attractive.

Now we prove the existence of an almost periodic solution. From almost periodicity of $y(t)$ (see Definition 2), it follows that there exists $l > 0$ such that any l -length interval contains a constant $\xi > 0$ so that $\|y(t + \xi) - y(t)\| \leq \epsilon$ holds for all $t \in \mathbb{R}$, namely ξ is an ϵ -translation number of $y(t)$. By the global attractivity of model (4) and the almost periodicity of $y(\cdot)$, it follows that for any constant $M > 0$, there exists $T > 0$ such that for any $t > T$,

$$\begin{aligned} & \|\psi(t + \xi, i, y) - \psi(t, i, y)\| \\ &= \|\psi(t, \psi(\xi, i, y), \theta(\xi, y)) - \psi(t, i, y)\| \\ &\leq \|\psi(t, \psi(\xi, i, y), \theta(\xi, y)) - \psi(t, i, \theta(\xi, y))\| \\ &\quad + \|\psi(t, i, \theta(\xi, y)) - \psi(t, i, y)\| \\ &\leq M\epsilon. \end{aligned}$$

That is, $\psi(t, i, y)$ is asymptotically almost periodic.

Taking $\{t_k\}_{k \in \mathbb{N}}$ such that

$$\sup_{t, i} |f(i, y(t)) - f(i, y(t + t_k))| < 1/k$$

and considering the solution $\psi(t + t_k, i, y)$ to $di(t)/dt = f(i, y(t + t_k))$, $\forall k \in \mathbb{N}$. According to Lemma 2 in [48], we know that there exists a sub-sequence of $\{\psi(t + t_k, i, y)\}_{k \in \mathbb{N}}$, denoted by $\psi(t + t_k, i, y)$ for not to introduce extra notations, that converges uniformly and its limit at $k \rightarrow \infty$, denoted by $\psi^*(t, i, y)$, is a solution to model (4). Since $\psi(t + t_k, i, y)$ is asymptotically almost periodic, i.e., $|\psi(t + t_k + \xi, i, y) - \psi(t + t_k, i, y)| \leq M\epsilon$ for sufficiently large k , it follows that the limit $\psi^*(t, i, y)$ is almost periodic in t and ξ is an $M\epsilon$ -translation number of $\psi^*(t, i, y)$. \square

APPENDIX I PROOF OF THEOREM 6

Proof. For each $v \in V$, $i_v(t)$ in model (4) satisfies

$$\begin{aligned} \frac{di_v(t)}{dt} &= -h_v(i(t), \beta_v(t)) \cdot i_v(t) \\ &\quad + g_v(i(t), \alpha_v(t), \Gamma(t)) \cdot (1 - i_v(t)) \\ &\geq -\bar{h}_v(\bar{\beta}_v)i_v(t) + g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma})(1 - i_v(t)) \\ &= -[\bar{h}_v(\bar{\beta}_v) + g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma})]i_v(t) + g_v(i_{\min}, \underline{\alpha}_v, \underline{\Gamma}) \end{aligned}$$

and

$$\begin{aligned} \frac{di_v(t)}{dt} &\leq -\underline{h}_v(\underline{\beta}_v)i_v(t) + g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma})(1 - i_v(t)) \\ &= -[\underline{h}_v(\underline{\beta}_v) + g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma})]i_v(t) + g_v(i_{\max}, \bar{\alpha}_v, \bar{\Gamma}). \end{aligned}$$

By the Gronwall inequality [49], we obtain the bounds of $i_v(t)$. This completes the proof. \square