Active Cyber Defense Dynamics Exhibiting Rich Phenomena

Ren Zheng[1] Wenlian Lu[1] *Shouhuai Xu*[2]

[1]Fudan University

[2]University of Texas at San Antonio

HotSoS 2015

Outline

Introduction

- There is an asymmetry between *reactive cyber defenses* and *cyber attacks*
  - The effect of attacks is amplified by a **function** of λ_1 (i.e., network effect), while the effect of defense is not.
  - λ_1 is the spectral radius of the attack-defense interaction structure ("overlay" networks in most cases)

Introduction

- There is an asymmetry between *reactive cyber defenses* and *cyber attacks*
  - The effect of attacks is amplified by a **function** of λ_1 (i.e., network effect), while the effect of defense is not.
  - λ_1 is the spectral radius of the attack-defense interaction structure ("overlay" networks in most cases)
- How to eliminate the asymmetry that is to the advantage of the attacker?

Introduction

- There is an asymmetry between *reactive cyber defenses* and *cyber attacks*
  - The effect of attacks is amplified by a **function** of λ_1 (i.e., network effect), while the effect of defense is not.
  - λ_1 is the spectral radius of the attack-defense interaction structure ("overlay" networks in most cases)
- How to eliminate the asymmetry that is to the advantage of the attacker?
- Active Cyber Defense (ACD) is one approach [Internet Mathematics 2015]

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)
- The idea of ACD is not new, but the rigorous characterization of its power is.

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)
- The idea of ACD is not new, but the rigorous characterization of its power is.
- Our goal: Understanding ACD’s power and limitation

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)
- The idea of ACD is not new, but the rigorous characterization of its power is.
- Our goal: Understanding ACD’s power and limitation
  - When is ACD effective? [Internet Mathematics 2015]

Active Cyber Defense (ACD)

- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)
- The idea of ACD is not new, but the rigorous characterization of its power is.
- Our goal: Understanding ACD’s power and limitation
  - When is ACD effective? [Internet Mathematics 2015]
  - How to use ACD optimally? [GameSec’13]

Active Cyber Defense (ACD)

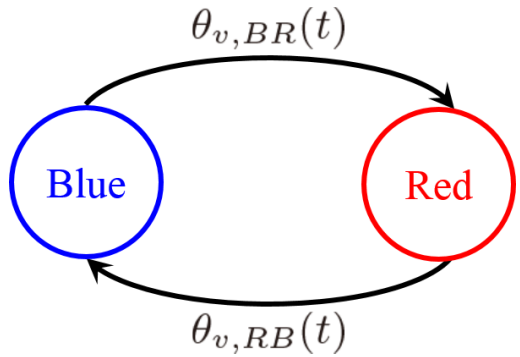
- Idea: Defender uses defenseware (e.g., “white worms”) to cure infected computers (under its jurisdiction)
- Disclaimer: ACD $\neq$ Fight Back (Retaliation)
- The idea of ACD is not new, but the rigorous characterization of its power is.
- Our goal: Understanding ACD’s power and limitation
  - When is ACD effective? [Internet Mathematics 2015]
  - How to use ACD optimally? [GameSec’13]
  - This paper: Rich phenomena that can be exhibited by ACD and their implications

Model

- Vertex set $V = \{1, 2, \cdots, n\}$ representing computers (or components)
- Attack-victim relation formulates an attack structure $G_R = (V, E_R)$, represented by adjacency matrix $A_R = [a^R_{vu}]_{n \times n}$
- ACD formulates a defense structure $G_B = (V, E_B)$, represented by adjacency matrix $A_B = [a^B_{vu}]_{n \times n}$
- Each node $v \in V$ has two possible states: secure or Blue; compromised or Red
- $B_v(t)$: the probability node $v \in V$ is in state Blue at time t
- $R_v(t)$: the probability node $v \in V$ is in state Red at time t

Model

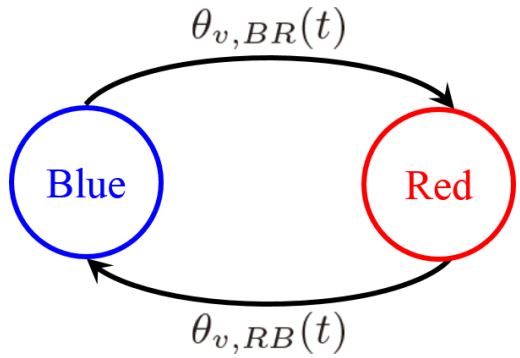
The state transition diagram for each node $v \in V$



$\theta_{V,BR}(t)$: probability v changes state from Blue to Red at time t,
$\theta_{V,RB}(t)$: probability v changes state from Red to Blue at time t

Model

The state transition diagram for each node $v \in V$



$\theta_{V,BR}(t)$: probability v changes state from Blue to Red at time t,
$\theta_{V,RB}(t)$: probability v changes state from Red to Blue at time t

This leads to the master equation for each $v \in V$:

$$\begin{cases} \dfrac{\mathrm{d}B_V(t)}{\mathrm{d}t} = \theta_{V,RB}(t) \cdot R_V(t) - \theta_{V,BR}(t) \cdot B_V(t) \\ \dfrac{\mathrm{d}R_V(t)}{\mathrm{d}t} = \theta_{V,BR}(t) \cdot B_V(t) - \theta_{V,RB}(t) \cdot R_V(t) \end{cases}, \tag{1}$$

How to specify $\theta_{v,BR}(t)$ and $\theta_{v,RB}(t)$?

$\theta_{v,BR}(t)$ depends on defense-power function $f : [0,1] \to \{0\} \cup \mathbb{R}^+$, which abstracts the power of the defenseware in detecting and cleaning up compromised (Red) nodes:

$$\theta_{v,RB}(t) = f\left(\frac{1}{\deg(v,G_B)} \sum_{u \in N_{v,G_B}} B_u(t)\right),$$

N_{v,G_B} is v's neighbors and $deg(v, G_B)$ is its (in)-degree in G_B.

How to specify $\theta_{v,BR}(t)$ and $\theta_{v,RB}(t)$?

$\theta_{v,BR}(t)$ depends on defense-power function $f : [0,1] \to \{0\} \cup \mathbb{R}^+$, which abstracts the power of the defenseware in detecting and cleaning up compromised (Red) nodes:

$$\theta_{v,RB}(t) = f\left(\frac{1}{\deg(v,G_B)} \sum_{u \in N_{v,G_B}} B_u(t)\right),$$

N_{v,G_B} is v's neighbors and $deg(v, G_B)$ is its (in)-degree in G_B.

Similarly, $\theta_{v,RB}(t)$ depends on attack-power function $g : [0,1] \to \{0\} \cup \mathbb{R}^+$, which abstracts the power of the attack (e.g., malware) in compromising secure (Blue) nodes:

$$\theta_{v,BR}(t) = g\left(\frac{1}{\deg(v,G_R)} \sum_{u \in N_{v,G_R}} B_u(t)\right),$$

Properties of defense- and attack-power functions

- Defense-power function $f(\cdot) \geq 0$
- $f(0) = 0$: ACD must be launched from some Blue node
- Attack-power function $g(\cdot) \geq 0$
- $g(1) = 0$: attack must be launched from some Red node
- The two functions, $f(\cdot)$ and $g(\cdot)$ do not have to have any specific relation, except that they are differentiable (for analytical treatment).

Main research task

For all $t \geq 0$ and all $v \in V$

- $\frac{\mathrm{d}B_v(t)}{\mathrm{d}t} + \frac{\mathrm{d}R_v(t)}{\mathrm{d}t} = 0$
- $B_v(t) + R_v(t) = 1$

Therefore, we only need to consider:

$$\begin{aligned} \frac{\mathrm{d}B_v(t)}{\mathrm{d}t} = \quad & f\left(\frac{1}{\deg(v,G_B)}\sum_{u\in N_{v,G_B}} B_u(t)\right)\left[1 - B_v(t)\right] \\ & -g\left(\frac{1}{\deg(v,G_R)}\sum_{u\in N_{v,G_R}} B_u(t)\right)B_v(t) \end{aligned} \tag{2}$$

Main research task is to analyze system (2) for all $v \in V$

Equilibrium is a useful concept for cyber security

- Despite that cyber security is rarely be in equilibrium
- We can quantify the effectiveness of ACD via the notion of σ-effectiveness: the dynamics converges to σ
- We consider homogeneous equilibria $[B_1^*, \cdots, B_n^*]$ with $B_1^* = \cdots = B_n^* = \sigma \in [0, 1]$
  - All-Blue equilibrium $B^* = \mathbf{1}$: ACD is 1-effective
  - All-Red equilibrium $B^* = \mathbf{0}$: ACD is 0-effective
  - σ-equilibrium $B^* = \sigma$: ACD is σ-effective
  - Stability captures a certain kind of resilience

A tool: Jacobian matrix

The Jacobian matrix of Eq. (2) near equilibrium σ is:

$$M = \left[(1-\sigma)f'(\sigma)D_{A_B}^{-1}A_B - \sigma g'(\sigma)D_{A_R}^{-1}A_R\right] - \left[f(\sigma)+g(\sigma)\right]I_n.$$

A tool: Jacobian matrix

The Jacobian matrix of Eq. (2) near equilibrium σ is:

$$M = \left[(1-\sigma)f'(\sigma)D_{A_B}^{-1}A_B - \sigma g'(\sigma)D_{A_R}^{-1}A_R\right] - \left[f(\sigma)+g(\sigma)\right]I_n.$$

Let $\lambda(M)$ be the set of eigenvalues of matrix M

A tool: Jacobian matrix

The Jacobian matrix of Eq. (2) near equilibrium σ is:

$$M = \left[(1-\sigma)f'(\sigma)D_{A_B}^{-1}A_B - \sigma g'(\sigma)D_{A_R}^{-1}A_R\right] - \left[f(\sigma)+g(\sigma)\right]I_n.$$

Let $\lambda(M)$ be the set of eigenvalues of matrix M

Hypothesis H_0: existence of homogeneous equilibrium in Eq. (2)

There exists $\sigma \in [0,1]$ such that $(1-\sigma)\,f(\sigma) = \sigma\, g(\sigma)$.

A tool: Jacobian matrix

The Jacobian matrix of Eq. (2) near equilibrium σ is:

$$M = \left[(1-\sigma)f'(\sigma)D_{A_B}^{-1}A_B - \sigma g'(\sigma)D_{A_R}^{-1}A_R\right] - \left[f(\sigma)+g(\sigma)\right]I_n.$$

Let $\lambda(M)$ be the set of eigenvalues of matrix M

Hypothesis H_0: existence of homogeneous equilibrium in Eq. (2)

There exists $\sigma \in [0,1]$ such that $(1-\sigma)\,f(\sigma) = \sigma\, g(\sigma)$.

This hypothesis will be replaced by actual conditions in the full version of the paper.

Existence and stability of equilibria

Proposition 1

Under hypothesis $\boldsymbol{H}_0$, equilibrium $B^* = \sigma \in [0, 1]$ of Eq. (2) is stable if the real part $\Re(\mu) < 0$ for all $\mu \in \lambda(M)$, and unstable if $\Re(\mu) > 0$ for some $\mu \in \lambda(M)$.

Corollary 1

If $f(0) = 0$, then equilibrium $B^* = 0$ is locally stable when $f'(0) < g(0)$ and locally unstable when $f'(0) > g(0)$.

If $g(1) = 0$, then equilibrium $B^* = 1$ is locally stable when $-g'(1) < f(1)$ and locally unstable when $-g'(1) > f(1)$.

Existence and stability of equilibria (cont.)

Corollary 2

Suppose $G_B = G_R = G$ (i.e., $A_B = A_R = A$). Let μ_1 be the eigenvalue of $D_A^{-1}A$ that has the smallest real part. If the attack-power and defense-power satisfy one of the following:

- $(1-\sigma)f'(\sigma) - \sigma g'(\sigma) > 0$ and $\dfrac{f(\sigma)+g(\sigma)}{(1-\sigma)f'(\sigma)-\sigma g'(\sigma)} > 1$
- $(1-\sigma)f'(\sigma) - \sigma g'(\sigma) < 0$ and $\dfrac{f(\sigma)+g(\sigma)}{(1-\sigma)f'(\sigma)-\sigma g'(\sigma)} < \Re(\mu_1)$,

then $B^* = \sigma \in [0,1]$ is locally stable. If the attack-power and defense-power satisfy one of the following:

- $(1-\sigma)f'(\sigma) - \sigma g'(\sigma) > 0$ and $\dfrac{f(\sigma)+g(\sigma)}{(1-\sigma)f'(\sigma)-\sigma g'(\sigma)} < 1$
- $(1-\sigma)f'(\sigma) - \sigma g'(\sigma) < 0$ and $\dfrac{f(\sigma)+g(\sigma)}{(1-\sigma)f'(\sigma)-\sigma g'(\sigma)} > \Re(\mu_1)$,

then $B^* = \sigma \in [0,1]$ is locally unstable.

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2$

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Scenario 2: $f(x) = x^2 + x$

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Scenario 2: $f(x) = x^2 + x \Rightarrow B^* = 0$ is unstable, $B^* = 1$ is stable.

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Scenario 2: $f(x) = x^2 + x \Rightarrow B^* = 0$ is unstable, $B^* = 1$ is stable.

Scenario 3: $f(x) = x^2 + \frac{1}{2}x$

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Scenario 2: $f(x) = x^2 + x \Rightarrow B^* = 0$ is unstable, $B^* = 1$ is stable.

Scenario 3: $f(x) = x^2 + \frac{1}{2}x \Rightarrow B^* = 0$ and $B^* = 1$ are stable, $B^* = 0.5$ is unstable.

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

Scenario 2: $f(x) = x^2 + x \Rightarrow B^* = 0$ is unstable, $B^* = 1$ is stable.

Scenario 3: $f(x) = x^2 + \frac{1}{2}x \Rightarrow B^* = 0$ and $B^* = 1$ are stable, $B^* = 0.5$ is unstable.

Scenario 4: $f(x) = -2x^2 + 2x$

Example: stability effect

Suppose $G_B = G_R$ is an Erdös-Rényi (ER) random graph instance $G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Consider attack-power function $g(x) = 1 - x$ against the following defense-power function $f(x)$:

Scenario 1: $f(x) = x^2 \Rightarrow B^* = 0$ is stable, $B^* = 1$ is unstable.

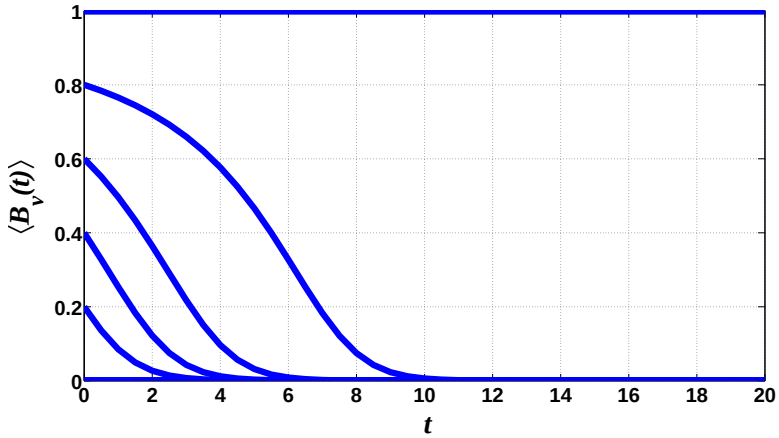
Scenario 2: $f(x) = x^2 + x \Rightarrow B^* = 0$ is unstable, $B^* = 1$ is stable.

Scenario 3: $f(x) = x^2 + \frac{1}{2}x \Rightarrow B^* = 0$ and $B^* = 1$ are stable, $B^* = 0.5$ is unstable.

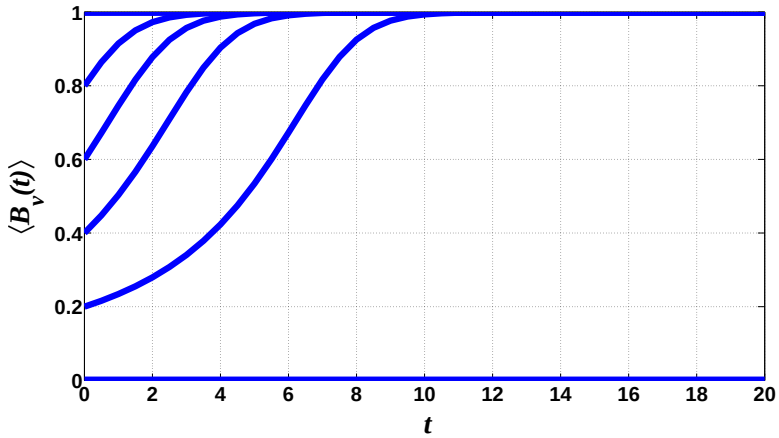
Scenario 4: $f(x) = -2x^2 + 2x \Rightarrow B^* = 0$ and $B^* = 1$ are unstable, $B^* = 0.5$ is stable.

Example: stability effect

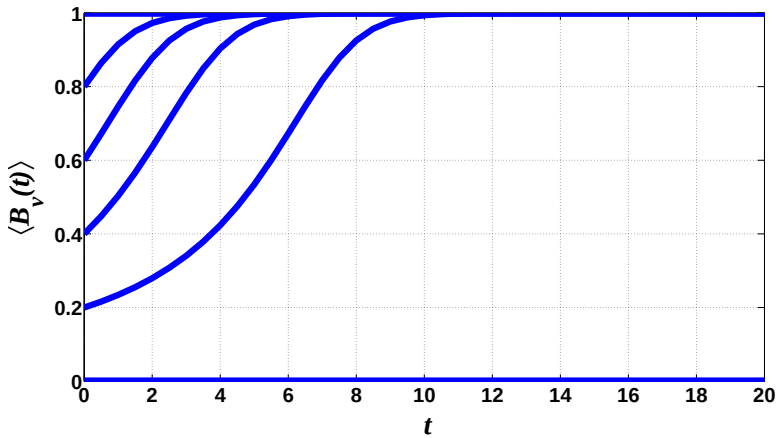
Metric: $\langle B_v(t) \rangle = \frac{1}{|V|} \sum_{v \in V} B_v(t)$, portion of secure nodes.



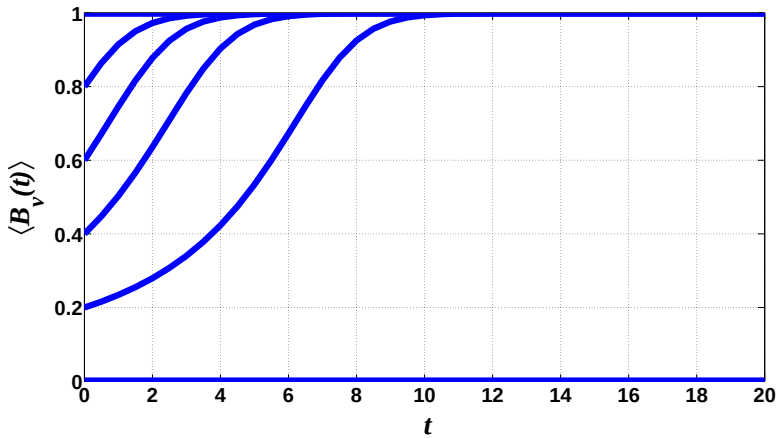
(a) Scenario I: $f(x) = x^2$



(b) Scenario II: $f(x) = x^2 + x$



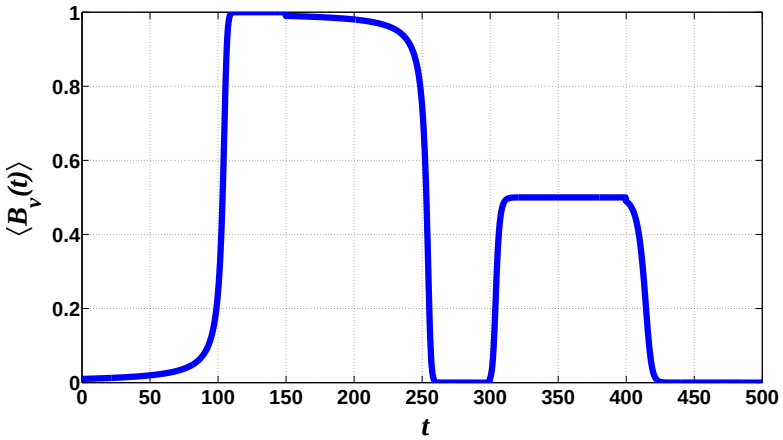
(c) Scenario III: $f(x) = x^2 + \frac{1}{2}x$



(d) Scenario IV: $f(x) = -2x^2 + 2x$

Why the concept of equilibrium is still useful even though cyber security may rarely be in equilibrium?

Time t	$f(x)$	$g(x)$	B^*
[0,150]	$f(x) = x^2 + x$	$g(x) = 1 - x$	$B^* = 1$
[150,300]	$f(x) = x^2$	$g(x) = 1 - x$	$B^* = 0$
[300,400]	$f(x) = -2x^2 + 2x$	$g(x) = 1 - x$	$B^* = 0.5$
[400,500]	$f(x) = x^2 + \frac{1}{2}x$	$g(x) = 1 - x$	$B^* = 1$



(e) Dynamic defense power $f(x)$ (f) This is why!!

Figure: Effects of perturbations at $t = 150, 300, 400$

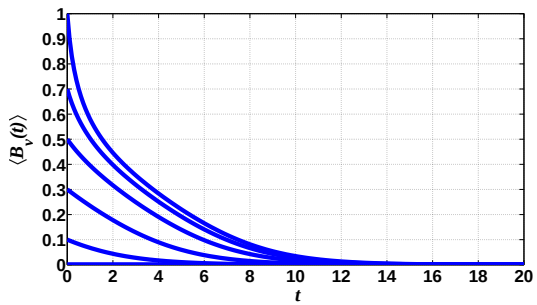
Insight

ACD dynamics may not enter into "equilibrium" because of perturbations to security state (e.g., cleaning/removing some compromised computers), and/or perturbations to attack/defense power (e.g., introduction of new attack/defense).

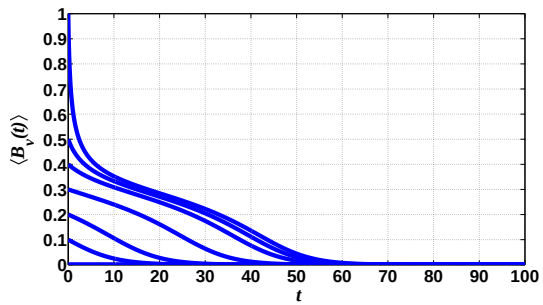
Warm up: parameterized defense-power

Suppose $G_B = G_R$ is ER graph instance $G = (V, E)$ with $|V| = 2,000$ and $p = 0.5$. Consider parameterized defense-power $f(x, \nu)$ with parameter $\nu \in (0, +\infty)$ and attack-power $g(x)$:

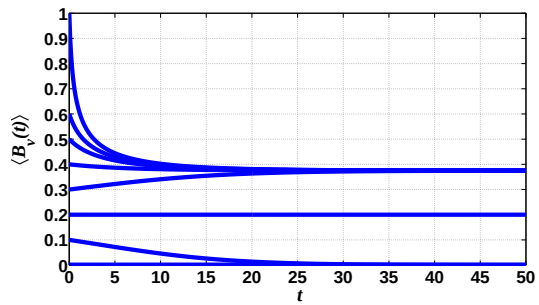
$$f(x, \nu) = \nu x - 2x^2, \quad g(x) = (1 - 2x)^2$$



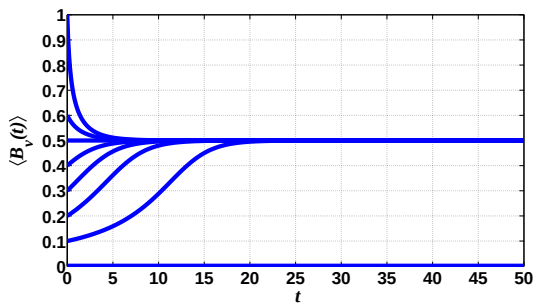
(a) $\nu = 0.5$



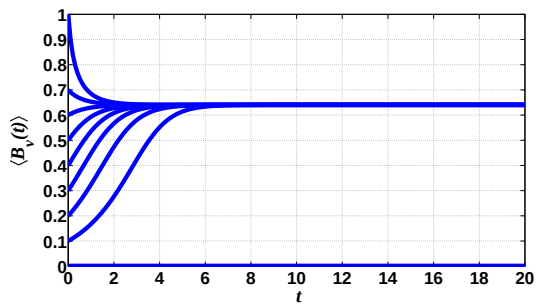
(b) $\nu = 0.8$



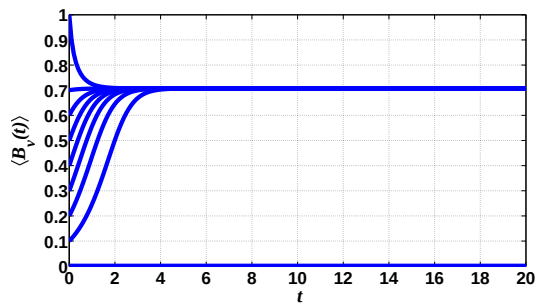
(c) $\nu = 0.85$



(d) $\nu = 1$



(e) $\nu = 1.5$



(f) $\nu = 2$

Phenomenon: Transition between multiple attractors

Under condition $f(0) = g(1) = 0$, system (2) has two locally stable equilibria $B^* = 1$ and $B^* = 0$.

Given thresholds $\tau_1^*, \tau_2^* \in (0, 1)$. Define Blue and Red threshold sets as:

$$\Xi_{G_B,\tau_1^*} = \left\{ B \in [0,1]^n \;\middle|\; \frac{1}{\deg(v, G_B)} \sum_{u \in N_{v,G_B}} B_u \geq \tau_1^*, \forall v \in V \right\}.$$

$$\Xi_{G_R,\tau_2^*} = \left\{ R \in [0,1]^n \;\middle|\; \frac{1}{\deg(v, G_R)} \sum_{u \in N_{v,G_R}} R_u \geq \tau_2^*, \forall v \in V \right\}.$$

Manipulating the initial state $B(0)$ can cause transition of the dynamics between the two equilibria!

Phenomenon: Transition between multiple attractors

Theorem 1

Let $G_B = (V, E_B)$ and $G_R = (V, E_R)$ be two arbitrary graphs. Suppose $f(\cdot)$ and $g(\cdot)$ are continuous with $f(0) = g(1) = 0$.

Case 1: Suppose attack-power and defense-power satisfy: for any $z \in [\tau_1^*, 1)$, $B \in \Xi_{G_B,\tau_1^*}$ and some $\alpha > 0$ we have $f(z) > \alpha \cdot z$ and

$$f\left(\frac{1}{\deg(v, G_B)} \sum_{u \in N_{v,G_B}} B_u\right) + g\left(\frac{1}{\deg(v, G_R)} \sum_{u \in N_{v,G_R}} B_u\right) \leq \alpha.$$

If initial value $B(0) \in \Xi_{G_B,\tau_1^*}$, $\lim_{t\to\infty} B_v(t) = 1$ for $v \in V$.

Case 2: Suppose for any $z \in [\tau_2^*, 1)$, $R \in \Xi_{G_R,\tau_2^*}$ and some $\beta > 0$ we have $g(1 - z) > \beta \cdot z$ and

$$f\left(1 - \frac{1}{\deg(v, G_B)} \sum_{u \in N_{v,G_B}} R_u\right) + g\left(1 - \frac{1}{\deg(v, G_R)} \sum_{u \in N_{v,G_R}} R_u\right) \leq \beta$$

If initial value $R(0) \in \Xi_{G_R,\tau_2^*}$, $\lim_{t\to\infty} R_v(t) = 1$ for $v \in V$.

Transition between multiple attractors

Cyber security meaning of the Theorem

Under a certain condition (**case 1**), the defender needs to manipulate the initial security state $\mathbf{B}(0)$ to belong to Ξ_{G_B,τ_1^*} to make active defense *1-effective*; this says what the defender should strive to do.

Transition between multiple attractors

Cyber security meaning of the Theorem

Under a certain condition (**case 1**), the defender needs to manipulate the initial security state $\mathbf{B}(0)$ to belong to Ξ_{G_B,τ_1^*} to make active defense *1-effective*; this says what the defender should strive to do.

Under certain other circumstances (**case 2**), the defender should make sure that the initial security state $\mathbf{B}(0)$ does not cause $\mathbf{R}(0) = \mathbf{1} - \mathbf{B}(0) \in \Xi_{G_R,\tau_2^*}$, because in this regime active defense is *0-effective*; this says what the defender should strive to avoid.

Example: Transition between $B^* = 0$ and $B^* = 1$

Consider defense-power $f(x) = \dfrac{1}{e^{-10x+5} + 1}$ and attack-power $g(x) = 2(1-x)^2$.

Example: Transition between $B^* = 0$ and $B^* = 1$

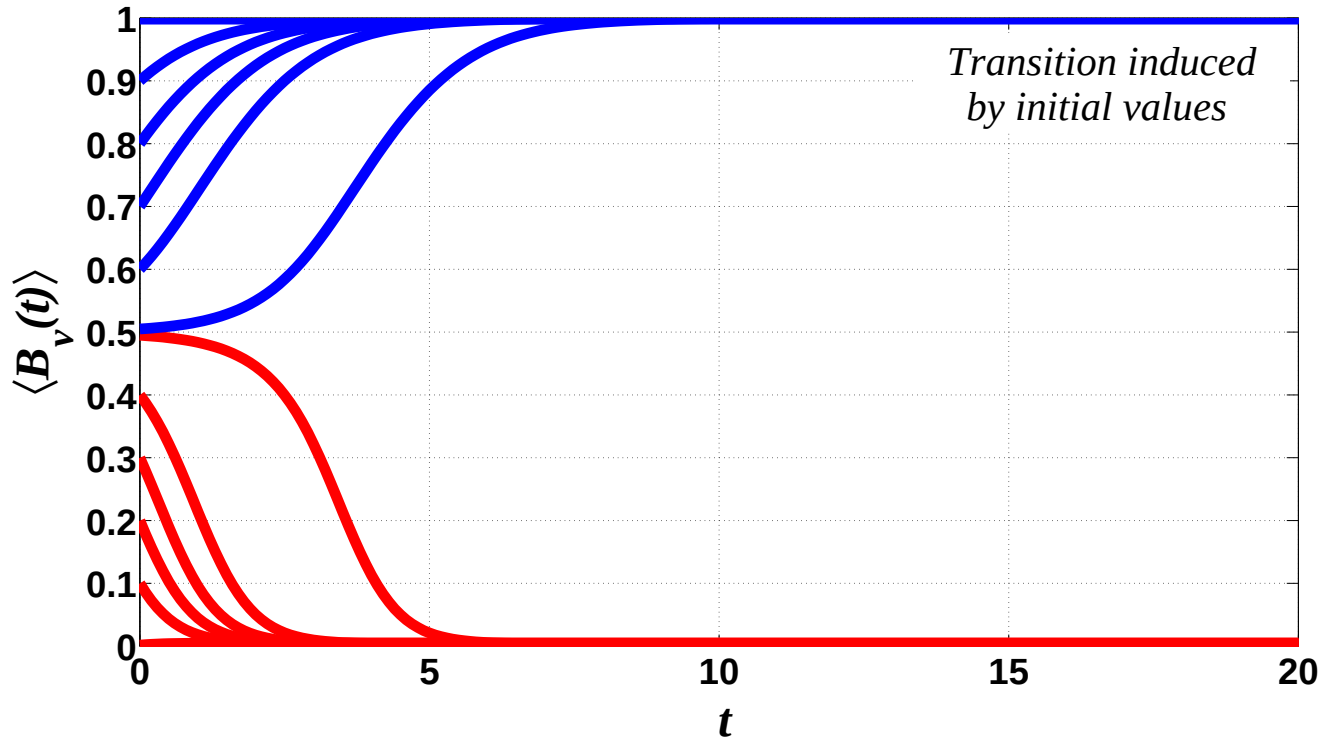
Consider defense-power $f(x) = \frac{1}{e^{-10x+5} + 1}$ and attack-power $g(x) = 2(1-x)^2$. G_B and G_R are two different ER graph instances with $|V| = 2,000$ and $p = 0.5$.

Example: Transition between $B^* = 0$ and $B^* = 1$

Consider defense-power $f(x) = \frac{1}{e^{-10x+5}+1}$ and attack-power $g(x) = 2(1-x)^2$. G_B and G_R are two different ER graph instances with $|V| = 2,000$ and $p = 0.5$. By manipulating the initial value $\langle B_v(0) \rangle$ to $\langle B_v(0) \rangle > 0.5$, the system converges to $B^* = 1$;

Example: Transition between $B^* = 0$ and $B^* = 1$

Consider defense-power $f(x) = \dfrac{1}{e^{-10x+5}+1}$ and attack-power $g(x) = 2(1-x)^2$. G_B and G_R are two different ER graph instances with $|V| = 2,000$ and $p = 0.5$. By manipulating the initial value $\langle B_v(0)\rangle$ to $\langle B_v(0)\rangle > 0.5$, the system converges to $B^* = 1$; By manipulating the initial value $\langle B_v(0)\rangle$ to $\langle B_v(0)\rangle < 0.5$, the system converges to $B^* = 0$.



Insight

A small change in the initial global security state, in the model parameters, in the attack network structure, or in the defense network structure can lead to substantial change in ACD dynamics.

Insight

A small change in the initial global security state, in the model parameters, in the attack network structure, or in the defense network structure can lead to substantial change in ACD dynamics. A rigorous characterization, such as the theorem mentioned above, can offer precise guidance on “what the defender should strive to do” and “what the defender should strive to avoid.”

Phenomenon: Hopf Bifurcation

Consider a system of equations with differentiable F

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x, \nu), \quad x \in \mathbb{R}^n.$$

Basic idea: a critical value ν^* at which rich phenomenon emerges

Technical issue: How to identify the critical value ν^*?

Example: Hopf bifurcation by perturbation to para.

Technical issue: How to identify the critical value ν^*?

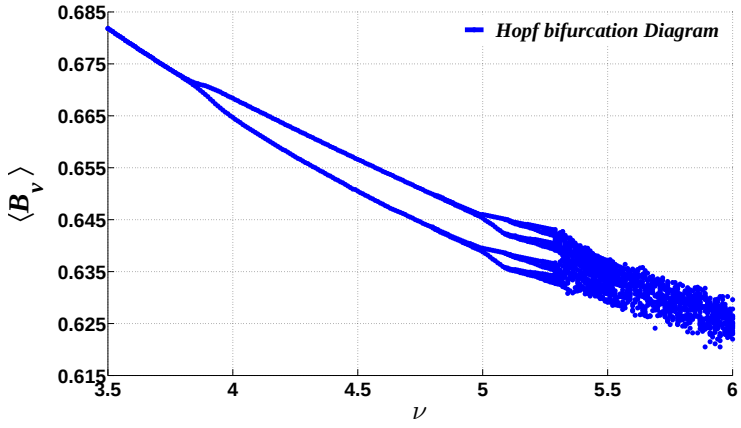
Consider ER graph instance $G_B = G_R = G = (V, E)$ with $|V| = 2,000$ and edge probability $p = 0.005$. Defense-power and attack-power functions are:

$$f(x) = -4x^2 + 4x, \quad g(x, \nu) = \left(\nu x - \frac{\nu}{2}\right)^2.$$

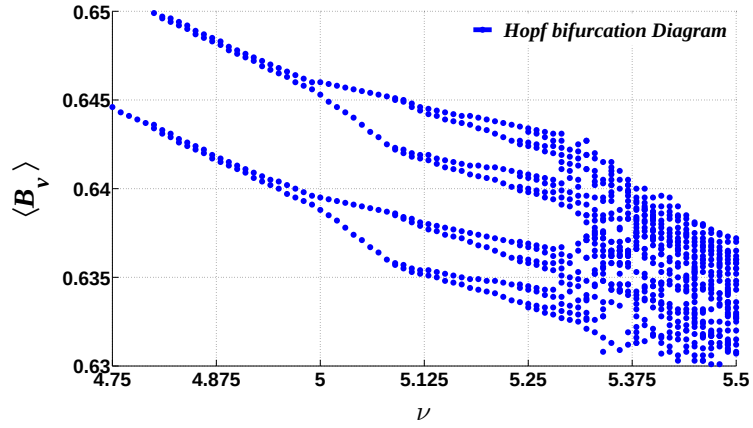
The real part of the eigenvalue of $D_A^{-1}A$ with the smallest real part is $\Re(\mu_1) = -0.3448$.

- When $\nu = 3$, we have $\Re(\mu_1) > -0.4$.
- When $\nu = 4$, we have $\Re(\mu_1) < -0.333$.
- When $\nu = 3.8$, we have $\Re(\mu_1) \approx -0.3448$.

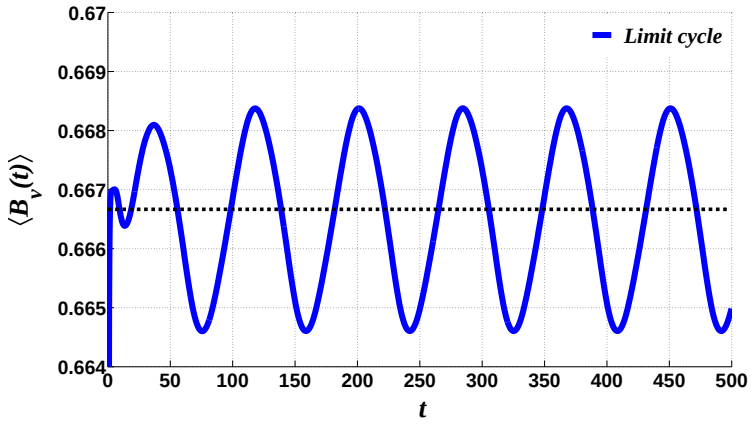
Example: Hopf bifurcation by perturbation to para.



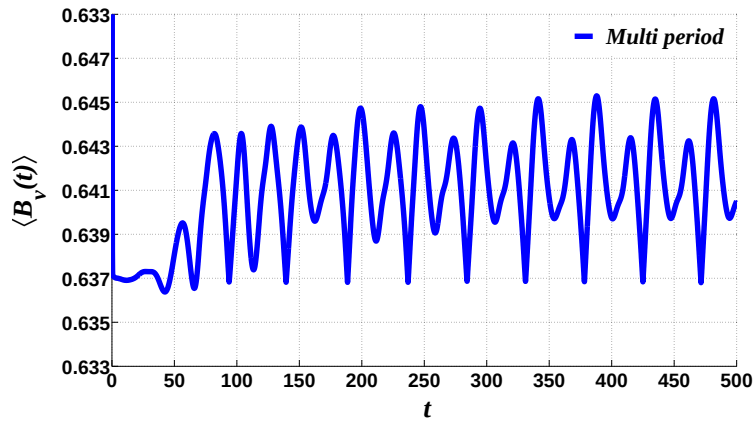
(k) Bifurcation diagram: $\nu \in (3, 6)$



(l) Zoon into: $\nu \in (4.75, 5.5)$



(m) Periodic trajectory: $\nu = 4$



(n) Periodic trajectory: $\nu = 5.05$

Figure: Hopf bifurcation diagram: $\langle B_V \rangle$ are the extremum points of $\langle B_V(t) \rangle$ in time period $t \in (1000, 2000)$.

Example: Hopf bifurcation by perturbation to G_R

Consider ER graph instances $G_B = (V, E_B)$ and $G_R = (V, E_R)$, both with $|V| = 2,000$ and $p = 0.005$. Consider defense-power and attack-power functions:

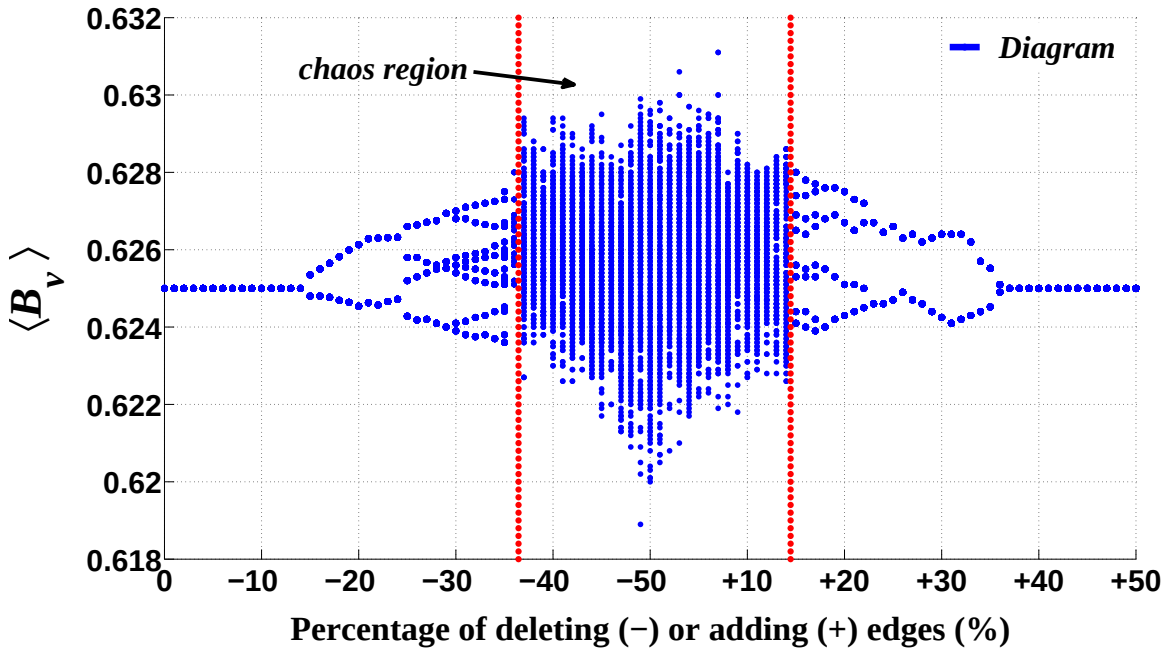
$$f(x) = -4x^2 + 4x, \quad g(x, \nu) = \left(\nu x - \frac{\nu}{2}\right)^2 \text{ with } \nu = 6$$

We make 100 iterations of perturbations to G_R as follows:

- During each of the first 50 iterations, we delete 226 edges (or 1% of the edges in the original E_R) chosen independently and uniformly at random
- During each of the next 50 iterations, we add 226 edges chosen independently and uniformly random among all the unconnected edges.

Example: Hopf bifurcation by perturbation to G_R

The *period-doubling cascade* phenomenon appears and finally leads to chaos after deleting $>$ 36% edges and before adding 14% edges. (Not symmetric because of randomness.)



Insight

ACD dynamics can exhibit Hopf bifurcation. These situations are "unmanageable:" infeasible/impossible to estimate the global security state in real-time, so we should avoid.

Phenomenon: Chaos

Let

$$F\left(B_v(t),t\right) = f\left(\frac{1}{\deg(v,G_B)}\sum_{u\in N_{v,G_B}} B_u(t)\right)\left[1-B_v(t)\right] - g\left(\frac{1}{\deg(v,G_R)}\sum_{u\in N_{v,G_R}} B_u(t)\right)B_v(t)$$

For a small perturbation to initial value $B_v(0)$, we have

$$\varepsilon_v(t) = D_x F(x,t)\Big|_{x=B_v(0)} \cdot \varepsilon_v(0)$$

where $\varepsilon_v(t)$ is a $n \times 1$ vector and $D_x F(x,t)\big|_{x=B_v(0)}$ is the Jacobian matrix of map F start at time t.

By the QR decomposition of matrix $\varepsilon(t) = [\varepsilon_1(t), \cdots, \varepsilon_n(t)]$, we obtain $\varepsilon(t) = q(t) \cdot r(t)$, where $q(t)$ is an orthogonal matrix and $r(t)$ is an upper triangular matrix.

Phenomenon: Chaos

To measure the average rate of convergence between two trajectories $F\left(B_v(0), t\right)$ and $F\left(B_v(0) + \varepsilon_v(0), t\right)$, we define:

Definition (Lyapunov characteristic exponents; LEs)

For a n-dimensional map, define Lyapunov characteristic exponents as

$$L_i = \lim_{t \to \infty} \frac{1}{t} \ln \left| \lambda_{ii}(t) \right| \quad (i = 1, \cdots, n),$$

$\lambda_{ii}(t)$ is the diagonal elements of upper triangular matrix $r(t)$.

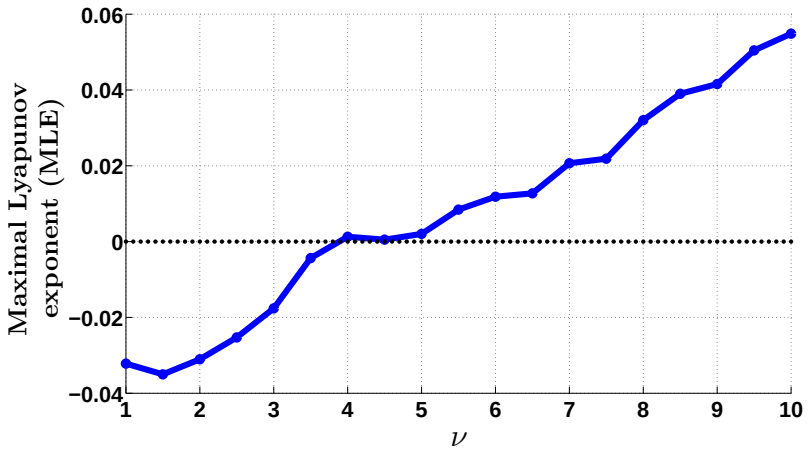
Under some mild contions, the limit exists for almost all initial values $B(0) = [B_1(0), \cdots, B_n(0)]$ and almost all matrix $\varepsilon(0)$.

The Maximal LE $MLE = \max_{1 \le i \le n} L_i$ indicates whether the dynamical system is chaotic (when $MLE > 0$) or not.

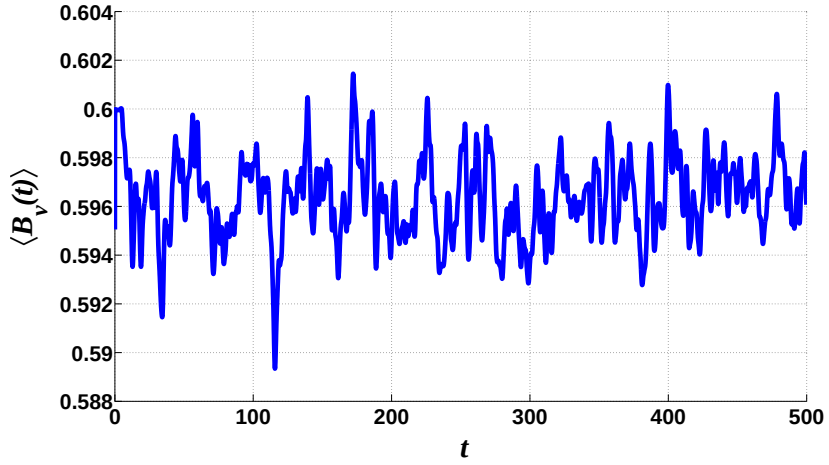
Example: Chaos

Consider an ER graph instance $G_B = G_R$ with $|V| = 2,000$ and $p = 0.005$, and defense-power and attack-power functions:

$$f(x) = -4x^2 + 4x \quad \text{and} \quad g_\nu(x) = \left(\nu x - \frac{\nu}{2}\right)^2.$$



(a) $MLE > 0$ indicates chaos



(b) $\langle B_\nu(t)\rangle$ for $\nu = 8$ exhibits chaos

Insight

ACD dynamics can be chaotic (infeasible/impossible to predict). Defender must manipulate the dynamics to avoid such "unmanageable" situations (e.g., by making $\nu \leq 5$ here).

Related work

- ACD is an integral component in the Cybersecurity Dynamics framework for modeling and quantifying cyber security from a holistic perspective
  - See [HotSoS'14] for history and numerous related works
- ACD is first *rigorously* modeled and studied by us in [Internet Mathematics'2015, GameSec'13]
- This paper: Understanding/characterizing ACD dynamics
  - Separate attack structure G_A from defense structure G_B
  - Consider attack-power and defense-power functions that cause ACD dynamics to exhibit rich phenomena
- Ongoing work: Earlywarning of such "critical transition" (without knowing the parameters)

Take-away message

- Birfucation and chaos are relevent to cyber security!
- Cyber security implications:
  - Infeasible/impossible to accurately measure/predict cyber security under certain circumstances
  - Defender must manipulate the dynamics to avoid such “un-manageable situations”
- This is an exciting field with many open problems!!