Characterizing the Power of Moving Target Defense via Cyber Epidemic Dynamics

Yujuan Han\textsuperscript{1,2}  Wenlian Lu\textsuperscript{1,3}  Shouhuai Xu\textsuperscript{2}

\textsuperscript{1}School of Mathematical Sciences  
Fudan University

\textsuperscript{2}Department of Computer Science  
University of Texas at San Antonio

\textsuperscript{3}Department of Computer Science  
University of Warwick

HotSoS 2014
Outline

Introduction

Cyber Epidemic Dynamics Model Accommodating MTD

Measuring the Power of MTD

MTD only induces dynamic parameters

MTD only induces dynamic attack-defense structures

Limitations of the Model

Related Work

Conclusion
Moving Target Defense (MTD) is believed to be a "game changer" for cyber defense.
Although there have been many MTD techniques, there is no systematic understanding and quantitative characterization of the power of MTD.
The power of MTD is often demonstrated via simulation.
Three classes of Moving Target Defense Techniques

- Networks-based MTD Techniques: IP address (and TCP port) randomization etc.
- Hosts-based MTD Techniques: instruction-level (ISR), code-level (code randomization), memory-level (ASLR), application-level (N-version programming, proactive cryptography)
- Instruments-based MTD Techniques: dynamic honeypot
Cyber Epidemic Dynamics Model

- Using *attack-defense* structure to capture the (attacker, victim) relation
- Using *parameters* to capture attacker and defense powers
- Using *epidemic threshold* to describe sufficient condition under which the epidemic dynamics converges to the equilibrium state — the clean state in this paper.
Our Contributions

- Idea: Using cyber epidemic dynamics to characterize the power of classes of MTD techniques (the lens)
- Techniques: Define and investigate two complementary measures (i.e., metrics).
  - Without considering cost of deploying MTD: the maximum portion of time (in the equilibrium) during which the system can stay in an undesired/insecure configuration/posture
  - Considering cost: the minimum cost of deploying MTD, while the system has to stay in an undesired/insecure configuration/posture for a given portion of time.
Cyber Epidemic Dynamics Model Accommodating MTD

MTD inducing dynamic attack-defense structures vs. MTD inducing dynamic parameters

- Networks-based MTD Techniques: *dynamic attack-defense structures*
- Hosts-based MTD Techniques: *dynamic parameters.*
- Instruments-based MTD Techniques: *dynamic attack-defense structures and dynamic parameters.*
Cyber Epidemic Dynamics Model

$G(t) = (V, E(t))$: attack-defense structure at time $t$, where $V$ is the vertex set (e.g., representing computers) and $E(t)$ is the edge-set representing the (attacker, victim) relation.

$A(t) = [A_{vu}(t)]$: adjacency matrix of $G(t)$

$i_v(t)$: the probability $v \in V$ is *infected* at time $t$

$\gamma(t)$: the probability an *infected* node $u \in V$ successfully attacks a *secure* node $v \in V$ over $(u, v) \in E(t)$ at time $t$

$\beta(t)$: the probability an *infected* node $v \in V$ becomes *secure* at time $t$

Suppose the attacks are independently launched, the model is,

$$
\frac{di_v(t)}{dt} = \xi_v(t)(1 - i_v(t)) - \beta(t)i_v(t)
$$

$$
= \left(1 - \prod_{u \in V} (1 - A_{vu}(t)i_u(t)\gamma(t))\right)(1 - i_v(t)) - i_v(t)\beta(t).
$$
Suppose both attack-defense structure and parameters are time-invariant \( t: G = (V, E) \) with adjacency matrix \( A; (\gamma, \beta) \).

The dynamics converges to equilibrium \( I^* = (0, \ldots, 0) \) if

\[
\mu \overset{\text{def}}{=} \beta - \gamma \lambda_1(A) > 0,
\]

where \( \lambda_1(A) \) is the largest (in modulus) eigenvalue of the adjacent matrix \( A \).

If \( \mu < 0 \), the dynamics does not converge to \( I^* = (0, \ldots, 0) \) at least for some initial values.
In this paper, MTD effectiveness is measured through the lens of occasionally deploying it to make the overall dynamics converge to $I^* = (0, \ldots, 0)$.

Suppose the defender can launch MTD to induce configurations $C_j = (G_j = (V, E_j), \beta_j, \gamma_j)$ for $2 \leq j \leq J$, which all satisfy condition (1).

If the defender always launch MTD to induce $C_j$, the problem is solved.

However, such MTD may not exist (e.g., attacker can introduce zero-day attacks) or the system to stay for some potion of time in undesired configuration $C_1$. 
Illustration of MTD-induced switching of configurations: The system is in configuration $C_1$ during time interval $[0, t_1)$, in $C_2$ during time interval $[t_1, t_2)$ because the defender launches MTD, etc. Although $C_1$ violates condition (1), the overall dynamics can converge to $I^* = (0, \ldots, 0)$ because of MTD. Note that $C_2$ and $C_3$ may reside in between two $C_1$’s (i.e., launching two combinations of MTD to induce $C_2$ and $C_3$ before returning to $C_1$).
Definition

$((\mu_1, \mu_2, \ldots, \mu_J, \pi^*_1)$-powerful MTD, without considering cost)

Denote by $\mu_k = \beta_k - \gamma_k \lambda_1(A_k)$ for $k = 1, \ldots, J$, where $A_k$ is the adjacency matrix of $G_k$.

1. Undesired configuration $C_1 = (G_1, \beta_1, \gamma_1)$ with $\mu_1 < 0$.
2. MTD induced configurations $C_j = (G_j, \beta_j, \gamma_j)$ with $\mu_j > 0$, $j \geq 2$.

We say MTD is $((\mu_1, \mu_2, \ldots, \mu_J, \pi^*_1)$-powerful if it can make the overall dynamics converge to $I^* = (0, \ldots, 0)$, while allowing the system to stay in configuration $C_1$ for the maximum $\pi^*_1$-portion of time in the equilibrium.
Measure II: Power of MTD while considering cost

Definition

\(((\mu_1, \mu_2, \cdots, \mu_J, \pi_1, \Upsilon))-\text{powerful, while considering cost}\)

Consider cost function \( h(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+ \) such that \( h(\mu_j) \) is the cost of launching MTD to induce configuration \( C_j \) for \( j = 2, \ldots, J \), where \( h'(\mu) \geq 0 \) for \( \mu > 0 \).

1. Undesired configuration \( C_1 = (G_1, \beta_1, \gamma_1) \), \( \mu_1 < 0 \), \( \pi_1 \) is the potion of time the system must stay in \( C_1 \).

2. MTD induced configurations \( C_j = (G_j, \beta_j, \gamma_j) \), \( \mu_j > 0 \), \( j \geq 2 \).

We say MTD is \( (\mu_1, \mu_2, \cdots, \mu_J, \pi_1, \Upsilon) \)-powerful if the overall dynamics converges to \( I^* = (0, \ldots, 0) \) at the minimum cost \( \Upsilon(\pi_2^*, \cdots, \pi_J^*) \), where \( \pi_j^* \) \( (2 \leq j \leq J) \) is the portion of time the system stays in configuration \( C_j \) in the equilibrium.
MTD only induces dynamic parameters

Theorem

(Xu et al., ACM TAAS 2014) Consider configurations
\((G, \beta(t), \gamma(t))\), where \((\beta(t), \gamma(t))\) are driven by a homogeneous
Markov process \(\eta_t\) with steady-state distribution \([\pi_1, \ldots, \pi_N]\)
and support \(\{(\beta_1, \gamma_1), \ldots, (\beta_N, \gamma_N)\}\), meaning
\[E(\beta_{\eta_t}) = \pi_1 \beta_1 + \cdots + \pi_N \beta_N \quad \text{and} \quad E(\gamma_{\eta_t}) = \pi_1 \gamma_1 + \cdots + \pi_N \gamma_N.\]
If
\[
\frac{\pi_1 \beta_1 + \cdots + \pi_N \beta_N}{\pi_1 \gamma_1 + \cdots + \pi_N \gamma_N} > \lambda_1(A),
\]
the dynamics will converge to \(I^* = (0, \ldots, 0)\); if
\[
\frac{\pi_1 \beta_1 + \cdots + \pi_N \beta_N}{\pi_1 \gamma_1 + \cdots + \pi_N \gamma_N} < \lambda_1(A),
\]
the dynamics will not converge to \(I^* = (0, \ldots, 0)\) at least for
some initial value scenarios.
Assumptions and Notations,

- Suppose the MTD only induces parameters \((\beta(t), \gamma(t))\), which are driven by a homogeneous Markov process \(\eta_t\), its steady-state distribution is \([\pi_1, \cdots, \pi_N]\) and support 
\[\{(\beta_1, \gamma_1), \cdots, (\beta_N, \gamma_N)\}\].
- Denote configurations \(C_j = (G_1, \beta_j, \gamma_j)\), \(\mu_j = \beta_j - \gamma_j \lambda_1(A_1)\) as \(\mu_1 < 0 < \mu_2 < \cdots < \mu_N\).
- Introduce a sufficiently constant \(0 < \delta \ll 1\) to make sure the MTD can reach the maximum portion of time in \(C_1\). (Since (1) is strict and the maximum value is always reached on the boundary).
MTD only induces dynamic parameters: power of MTD without Considering Cost

Theorem

For configurations $C_j = (G, \beta_j, \gamma_j)$ with $1 \leq j \leq N$, we have $\mu_j = \beta_j - \gamma_j \lambda_1(A)$ where $\mu_1 < 0 < \mu_2 < \cdots < \mu_N$. The maximal potion of time the system can afford to stay in configuration $C_1$ is

$$\pi_1^* = \frac{\mu_N - \delta}{\mu_N - \mu_1},$$

which is reached by launching MTD to induce $C_N$ only with portions of time given by

$$\pi_2^* = \cdots = \pi_{N-1}^* = 0, \quad \pi_N^* = \frac{\delta - \mu_1}{\mu_N - \mu_1}. \quad (2)$$

In other words, MTD is $(\mu_1, \cdots, \mu_N, \pi_1^*)$-powerful.
Algorithm for optimal MTD without considering cost

1. Compute $\pi_1^*$ according to (2).
2. **while** TRUE **do**
3. Wait for time $T_1 \leftarrow \exp(a/\pi_1^*)$ \{system in $C_1$\}
4. Launch MTD to make system stay in $C_N$ for time $T_N \leftarrow \exp(a/(1 - \pi_1^*))$
5. Stop launching MTD \{system returns to $C_1$\}
6. **end while**

$T \leftarrow \exp(e)$: assigning $T$ a value according to the exponential distribution with parameter $e$.
$a$: time resolution.
Dependence of $\pi_1^*$ on $-\mu_1$ and $\mu_N$. 
Optimal MTD while Considering Cost

Suppose $\pi_1$ is the potion of time the system must stay in $C_1$, it should satisfy $0 < \pi_1 \leq \frac{\mu_N - \delta}{\mu_N - \mu_1}$. $f(\cdot)$ is the cost function. The cost of launching MTD is

$$
\Phi(\pi_2, \ldots, \pi_N) = \pi_1 f(\mu_1) + \sum_{j=2}^{N} \pi_j f(\mu_j).
$$

Define

$$
\mu_{k^*} = \min \left\{ \mu_k | \mu_k > \frac{-\pi_1 \mu_1}{1 - \pi_1}, \ 2 \leq k \leq N \right\}
$$

(3)

and for $2 \leq l < m \leq N$,

$$
F(\mu_l, \mu_m) = \pi_1 f(\mu_1) + \frac{f(\mu_m) - f(\mu_l)}{\mu_m - \mu_l} (\delta - \pi_1 \mu_1) + \frac{\mu_m f(\mu_l) - \mu_l f(\mu_m)}{\mu_m - \mu_l} (1 - \pi_1).
$$

(4)
Theorem

If \( k^* = 2 \), the minimal cost is

\[
\min_{\pi_2, \ldots, \pi_N} \Phi(\pi_2, \ldots, \pi_N) = \pi_1 f(\mu_1) + (1 - \pi_1) f(\mu_2),
\]

which is reached by launching MTD to induce configuration \( C_2 \) only. If \( k^* > 2 \), the minimal cost is

\[
\min_{\pi_2, \ldots, \pi_N} \Phi(\pi_2, \ldots, \pi_N) = \min_{1 < k^* \leq m} F(\mu_l, \mu_m) = F(\mu_{l^*}, \mu_{m^*}).
\]  \hspace{1cm} (5)

The minimal cost is reached by launching MTD to induce configurations \( C_{l^*}, C_{m^*} \) respectively with portions of time:

\[
\begin{bmatrix}
\pi_{l^*} \\
\pi_{m^*}
\end{bmatrix} = \frac{1}{\mu_m^* - \mu_{l^*}} \begin{bmatrix}
(\mu_m^* - \delta) + \pi_1(\mu_1 - \mu_m^*) \\
-(\mu_{l^*} - \delta) + \pi_1(\mu_{l^*} - \mu_1)
\end{bmatrix}.
\]  \hspace{1cm} (6)

where \( 0 < \delta \ll 1 \) is some constant.
Algorithm for optimal MTD while considering cost

1. Compute $k^*$ according to (3)
2. **If** $k^* = 2$, wait in $C_1$ for time $T_1 \leftarrow \exp(a/\pi_1)$ and launch MTD to stay in $C_2$ for time $T_2 \leftarrow \exp(a/\pi_2)$ alternately.
3. **else** compute $\mu_1^*, \mu_m^*$ & $\pi_1^*, \pi_m^*$ according to (5)-(6). **endif**
4. Wait for time $T_1 \leftarrow \exp(a/\pi_1)$ \{system in $C_1$\}
5. Set $\Delta = \{l^*, m^*\}$, $j \leftarrow R \Delta$,
6. $T_j \leftarrow \exp(a/\pi_j)$.
7. Launch MTD to stay in $C_j$ for $T_j$.
8. Set $\Delta = \{1, l^*, m^*\} - \{j\}$, $j \leftarrow R \Delta$.
9. **If** $j = 1$, wait in $C_1$ for time $T_1 \leftarrow \exp(a/\pi_1)$, $j \leftarrow R \{l^*, m^*\}$, **endif**
If \( f(\cdot) \) is convex, the minimum cost of MTD is \( F(\mu_{k^*-1}, \mu_{k^*}) \), hence, the defender only needs to launch MTD to induce configurations \( C_{k^*-1}, C_{k^*} \).

- The complexity of searching for the optimal solution (i.e., \( k^* \) in this case) reduced to \( O(N) \).

If \( f(\cdot) \) is concave, the minimum cost of MTD is \( F(\mu_2, \mu_N) \), hence, the defender only needs to launch MTD to induce configurations \( C_2, C_N \).

- The complexity of searching for the optimal solution reduced to \( O(1) \).
MTD only induces dynamic attack-defense structures

Theorem
(a general result) Consider $C_l = (G_l, \beta, \gamma)$, $l = 1, \cdots, N'$, where $C_l = (G_l, \beta, \gamma)$ for $1 \leq l \leq j$ violate condition (1) but $C_k = (G_k, \beta, \gamma)$ for $j < k \leq N'$ satisfy condition (1). Then, MTD is effective if $G(t)$ are driven by Markov process strategy $\sigma_t$ with infinitesimal generator $Q = (q_{uv})_{N' \times N'}$ defined as:

(i) for $k > j$, $-q_{kk} \leq \frac{2a[\beta - \gamma \lambda_1(A_k) - \delta]}{j(c+N'-1-j-a)}$;

(ii) for $\ell \leq j$, $-q_{\ell\ell} \geq \frac{2b[\gamma \lambda_1(A_\ell) - \beta + \delta]}{b - \frac{c(j-1)}{N'-1} - \frac{N'-j}{N'-1}}$;

(iii) $q_{rp} = \frac{-q_{rr}}{N'-1}$ for all $p \neq r$ and $p, r \in \{1, \ldots, N'\}$.

here $0 < \delta \ll 1$, $c$ is related to the convergent speed, $a, b, c$ are arbitrary constants with $a < 1 < b < c$. 

MTD only induces dynamic attack-defense structures

- Suppose the MTD only induces parameters $G(t)$, which are driven by a homogeneous Markov process $\sigma_t$, its steady-state distribution is $[\pi_1, \cdots, \pi_{N'}]$ and support $\{G_1, \cdots, G_{N'}\}$.

- Denote configurations $C_j = (G_j, \beta_1, \gamma_1)$, $\mu_j = \beta_1 - \gamma_1 \lambda_1(A_j)$ as $\mu_1 < 0 < \mu_2 < \cdots < \mu_{N'}$.

- Introduce a sufficiently constant $0 < \delta \ll 1$ to make sure the MTD can reach the maximum portion of time in $C_1$. (Because of (1) is strict and the maximum value is always reached on the boundary).
MTD only induces dynamic attack-defense structures: power of MTD without Considering Cost

Theorem
For configurations $C_j = (G_j, \beta_1, \gamma_1)$ with $1 \leq j \leq N'$, we have $\mu_j = \beta_1 - \gamma_1 \lambda_1(A_j)$ and $\mu_1 < 0 < \mu_2 < \cdots < \mu_{N'}$. The maximal potion of time the system can afford to stay in configuration $C_1$ is

$$\pi_1^* = \frac{\frac{b-1}{2b[-\mu_1+\delta]}}{\frac{b-1}{2b[-\mu_1+\delta]} + \frac{c-a}{2a[\mu_{N'}-\delta]}},$$

(7)

where $0 < \delta \ll 1$, $a < 1 < b < c$, which is reached by launching MTD to induce $C_{N'}$ only with portion of time given by

$$\pi_{N'}^* = 1 - \pi_1^*.$$
Algorithm for optimal MTD without considering cost

1. Compute $\pi_1^*$ according to (7).
2. while TRUE do
3.   Wait for time $T_1 \leftarrow \exp(a/\pi_1^*)$ \{system in $C_1$\}
4.   Launch MTD to make system stay in $C_{N'}$ for time $T_{N'} \leftarrow \exp(a/(1 - \pi_1^*))$
5. Stop launching MTD \{system returns to $C_1$\}
6. end while
Dependence of $\pi_1^*$ on $-\mu_1$ and $\mu_{N'}$. 

Example Scenario
Steps to find minimum cost of launching MTD,

1. Consider possible combinations of MTD-induced configurations: \( \mathcal{L}_1, \ldots, \mathcal{L}_{2^{N'-1}} \)

2. Find \( \sigma_t \) (according to previous theorem) such that MTD forces the convergence to \( I^* = (0, \ldots, 0) \)

3. For each \( \mathcal{L}_i \) with valid \( \sigma_t \), find the maximum portion of time, denoted by \( \pi_i \), the MTD allows the system to stay in \( C_1 \). If \( \pi_i \geq \pi_1 \), keep \( \mathcal{L}_i \); otherwise, eliminate \( \mathcal{L}_i \).

4. For the remaining \( \mathcal{L}_j \)'s, compute the minimum cost of launching MTD corresponding to it.

5. Find the minimum cost among the costs.
Suppose $\pi_1$, where $\pi_1 \leq \pi_1^*$, is the potion of time the system must stay in $C_1$ and $g(\cdot)$ is the cost function. $\sigma_t$ defines the deployment of MTD: denote $Q = [q_{jk}]$ its infinitesimal generator, $x_l = \frac{1}{\sum_j \frac{1}{q_{lj}}}$ the expectation of sojourn time in $C_l$. Then, the portion of time in $C_l$ is $\pi_l = \frac{x_l}{\sum_j x_j}$.

Suppose MTD induces $C_{k_1}, \cdots, C_{k_{m'}}$, the cost of this MTD is

$$\Phi(\pi_2, \cdots, \pi_N) = \pi_1 g(\mu_1) + \sum_{j=2}^{N} \pi_j g(\mu_j)$$

$$= \pi_1 g(\mu_1) + (1 - \pi_1^*) \frac{\sum_{l=1}^{m'} x_{k_l} g(\mu_{k_l})}{\sum_{l=1}^{m'} x_{k_l}}.$$
Denote by

\[ G(k_1, \ldots, k_{m'}) = \frac{\sum_{l=1}^{\ell} \bar{x}_{k_l}(m') g(\mu_{k_l}) + g(\mu_{k_1}) \Delta(k_1, \ldots, k_{m'})}{\sum_{l=1}^{m'} \bar{x}_{k_l}(m') + \Delta(k_1, \ldots, k_{m'})}, \]

where

\[ \bar{x}_1 = \frac{b - 1}{2b[-\mu_1 + \delta]}, \quad \bar{x}_{k_l}(m') = \frac{c + m' - 1}{2a[\mu_{k_l} - \delta]}, \]

\[ \Delta(k_1, \ldots, k_{m'}) = \frac{1 - \pi_1}{\pi_1} \bar{x}_1 - \sum_{l=1}^{m'} \bar{x}_{k_l}(m'). \]

\[ \mathcal{K} = \left\{ \{k_1, \ldots, k_{m'}\} \mid \pi_1^* \leq \frac{\bar{x}_1}{\bar{x}_1 + \sum_{l=1}^{m'} \bar{x}_{k_l}(m')}, k_1 < \cdots < k_{m'} \right\}. \]
Optimal MTD while Considering Cost

**Theorem**

Find \( \{k_1^*, \cdots, k_m^*\} \) such that

\[
\{\mu_{k_1^*}, \cdots, \mu_{k_m^*}\} = \arg \min_{\{k_1, \cdots, k_m\} \in \mathcal{K}} G(k_1, \cdots, k_m)
\]  

(8)

For given cost function \( g(\cdot) \), the minimum cost is

\[
\Psi(\bar{x}_1, \bar{x}_{k_1^*}(m) + \Delta, \cdots, \bar{x}_{k_m^*}(m)) = \pi_1 g(\mu_1) + (1 - \pi_1) G(k_1^*, \cdots, k_m^*),
\]

which is reached by launching MTD to induce configuration \( \{(G_{k_i^*}, \beta, \gamma)\}_{i=1}^m \) via the following deployment strategy:

\[
\pi_{k_1^*} = (1 - \pi_1) \frac{\bar{x}_{k_1^*}(m) + \Delta(k_1^*, \cdots, k_m^*)}{\sum_{l=1}^m \bar{x}_{k_l^*}(m) + \Delta(k_1^*, \cdots, k_m^*)},
\]

(9)

\[
\pi_{k_l^*} = (1 - \pi_1) \frac{\bar{x}_{k_l^*}(m)}{\sum_{l=1}^m \bar{x}_{k_l^*}(m) + \Delta(k_1^*, \cdots, k_m^*)}, l = 2, \cdots, m.
\]
Algorithm for optimal MTD while considering cost

1. Compute $k_1^*, \ldots, k_m^*$ and $\pi_{k_1^*}, \ldots, \pi_{k_m^*}$ according to (8)-(9)
2. Wait for time $T_1 \leftarrow \exp(a/\pi_1)$ \{system in $C_1$\}
3. Set $\Delta = \{k_1^*, \ldots, k_m^*\}$, $k_j^* \leftarrow R \Delta$
4. $T_{k_j^*} \leftarrow \exp(a/\pi_{k_j^*})$
5. Launch MTD to stay in $C_{k_j^*}$ for time $T_{k_j^*}$
6. Set $\Delta = \{1, k_1^*, \ldots, k_m^*\} - \{k_j^*\}$, $k_j^* \leftarrow R \Delta$
7. If $k_j^* = 1$, wait in $C_1$ for time $T_1 \leftarrow \exp(a/\pi_1)$, $k_j^* \leftarrow R \{k_1^*, \ldots, k_m^*\}$ \textbf{endif}
8. Go to step 4
Limitations of the Model

1. Assume attack-defense structures and parameters are given.
2. Attacker cannot choose when to impose configuration $C_1$.
3. Assume homogeneous parameters $\gamma(v, u) = \gamma$ and $\beta(v) = \beta$. 
Characterizing effectiveness of MTD: two complementary perspectives:

- specific technique with localized view vs.
- classes of techniques with global view

Cyber Epidemic Dynamics: an active research area rooted in biological epidemic dynamics

- Our model is based on such model
We have introduced an approach of using cyber epidemic dynamics to characterize the power of MTD. The approach offers algorithms for optimally deploying MTD, where "optimization" means maximizing the portion of time the system can afford to stay in an undesired configuration, or minimizing the cost of launching MTD when the system has to stay in an undesired configuration for a predetermined portion of time.