FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Managing risks in large scale interdependent CPS.

(based on joint work with Aron Laszka and Shankar Shastry)

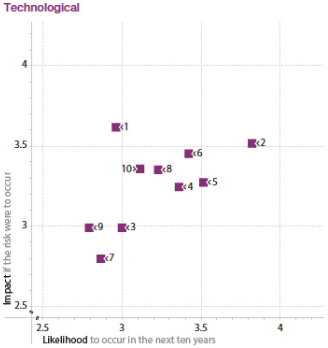
Galina Schwartz
Dept. of Electrical Engineering & Computer Sciences,
UC Berkeley, CA, USA

Berkeley University of California
Massachusetts Institute of Technology
NSF
Michigan
Vanderbilt University

Global interconnectedness ⟶ IDS nature of CPS risks

Cyber Risks: the Findings (based on Global Risk Reports)

- MIT Forum and Infosys Risk Group, survey based MIT Global Risk Survey, 06-2016
  - 92.54 percent of companies: the nature of risk is changing [due to complexity in the digital economy]
- World Economic Forum [WEF], expert based World Economic Forum, Global Risk Reports, yearly
  - Technology: highly varied expert opinions illustrated on the next slide



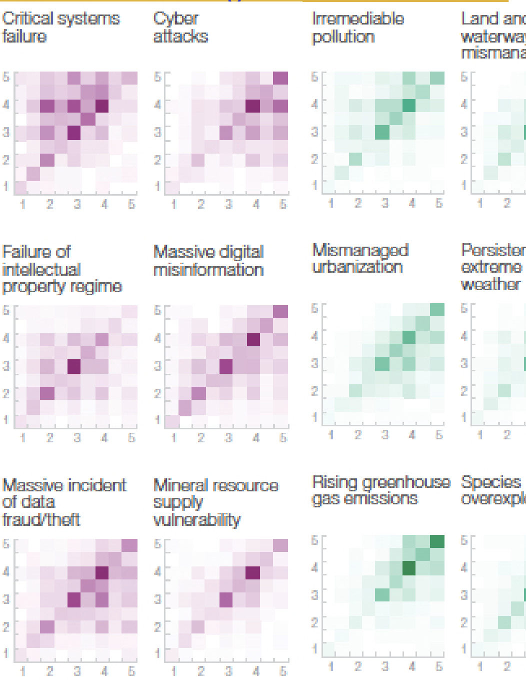
1	Critical systems failure	Single-point system vulnerabilities trigger cascading failure of critical information infrastructure and networks.
2	Cyber attacks	State-sponsored, state-affiliated, criminal or terrorist cyber attacks.
3	Failure of intellectual property regime	The loss of the international intellectual property regime as an effective system for stimulating innovation and investment.
4	Massive digital misinformation	Deliberately provocative, misleading or incomplete information disseminates rapidly and extensively with dangerous consequences.
5	Massive incident of data fraud/theft	Criminal or wrongful exploitation of private data on an unprecedented scale.
6	Mineral resource supply vulnerability	Growing dependence of industries on minerals that are not widely sourced with long extraction-to-market time lag for new sources.
7	Proliferation of orbital debris	Rapidly accumulating debris in high-traffic geocentric orbits jeopardizes critical satellite infrastructure.
8	Unforeseen consequences of climate change mitigation	Attempts at geoengineering or renewable energy development result in new complex challenges.
9	Unforeseen consequences of nanotechnology	The manipulation of matter on an atomic and molecular level raises concerns on nanomaterial toxicity.
10	Unforeseen consequences of new life science technologies	Advances in genetics and synthetic biology produce unintended consequences, mishaps or are used as weapons.

CPS risks assessment in the era of internet of things I

Pending questions

- How to measure?
- How to quantify?
- How to manage?

- At present:
  - cyber risks assessment is based on expert opinions
  - data is scarce
- Our task:
  - to develop sound valuation of CPS risks (statistics)
  - to take into account strategic nature of attacks (game theory)



Plan of the talk

- IDS: the main idea of the approach
- IDS model with discreet security choice
  - 2 player game
  - nonatomic players: identical and differing by security costs
  - Results:
    - Multiple equilibria could exist.
    - Present the tools of steering the system to superior equilibrium.
- IDS model with continuous security choice
  - atomic and non-atomic games
  - strategic attackers and defenders
  - endogenous player types (players choose their types)
  - Results:
    - Individually optimal security (Nash) differs from social optimum
    - Suggest the tools to shrink the inefficiency
- Novelty: we model IDS in large scale networks with strategic players
  - player choices are continuous
  - large scale IDS risks
  - strategic defenders
  - strategic attackers
  - network topology

Motivating Example: Attacks on electronic road signs

Dallas, TX, Interstate 30: Memorial day highway pranks

Saturday [May 28, 2016]

Tuesday morning [May 31, 2016]

Multiple attacks across USA.

http://www.worldwideinterweb.com/4812-funniest-hacked-traffic-signs/

My talk: Risk evaluation and management with interdependent security [IDS]

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

DOTs are shifting to electronic road signs

Texas Department of Transportation [TxDOT]

Dynamic messaging signs [DMS] "reduce confusion and increase safety"

From theory To practice

Tweaking of safety messages could lead to injuries or even deaths on the road. ... Third-degree felony (min 2 year sentence)

[TxDOT spokesman] Source: http://abc13.com/news/hackers-leave-quirky-messages-on-road-signs/1364333/

How to evaluate and manage the risks?

History of DMS (in)security: hacking remotely

Instructions: hacking DMS made by Daktronics [By SunHacker (2014)]

Sources: Security News, Brian Krebs Security Blog, Center for Internet Security (CIS) on malicious targeting of DMS

- Change the lan of VPN to INTERNET protocol
- Scan all the range of the IP on port 23
- Bruteforce the password (download scripts)
- Access the control panel; add your message

DHS alert: All Daktronics DMS

- Have the same default password
- Allow remote access to the control panel

Remote access (to control panel) ⟺
attacks may propagate indirectly [brave new world]

Electronic highway signs: hacking manually

Signs secured by (Buyers Barricades) [05-28-2016]. Turned off & locked (no remote access). TxDOT spokesman: Bold hacker(s) needed to:

- Power up the signs
- Break the password
- Manually alter the message via the control panel Source: http://www.techworm.net/2016/05/hacked-road-sign-texas-highway-says-trump-shape-shifting-lizard.html

No remote access ⟺
Indirect attacks are impossible [old world (no network effects)]

Prob. of breach with interdependent security [IDS]

No remote access $\Longleftrightarrow$ no indirect attacks $q = 0$ [old world]

Remote access $\Longleftrightarrow$ indirect attacks $q > 0$ [IDS]

- prob. of breach B [basic IDS], Kunreuther & Heal [2003], Hofmann [2007,2011]

$$B = P(d) + P(i) - P(d \cap i) = p + q(x) - pq(x) = p + (1-p)q(x)$$

  - p - prob. of direct loss; $q > 0$ - prob. of indirect loss $\Longleftarrow$ **important**
  - x fraction of insecure nodes, $q(x) > 0, q'(x) > 0, q(0) = 0, q(1) = \bar{q} < 1$

- prob. of breach [IDS for large systems] Öğüt at. al. [2005], Schwartz & Sastry [2014]

Independent nodes

$$B_i(p_1, ..., p_n) = 1 - s_i \prod_{j \neq i}^{n} (1 - (1 - s_j) q_{ij}), \quad s_i = 1 - p_i$$

  - p_i - prob. of direct loss; $q_{ij} \geq 0$ - prob. of indirect attack from node j to i

Further motivating examples: infrastructure CPS and Internet of things

Smart, Networked, Interconnected = IDS [$q > 0$]

- Electric grid
  - smart meters reprogramming
  - remote alteration of customer records
- Auto safety trade-offs: remote updates $\longrightarrow$ remote exploits
  - car owner: altering engine electronics (improved performance, higher emissions)
  - extortion of a car owner (via hacking smart auto software)
- Connections between infrastructures: (ex. Nest thermostat)

  http://www.tomsguide.com/us/nest-weave-smart-home,news-21658.html

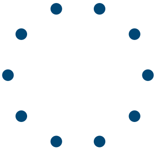
The size of q reflects

- network topology
- degree of interdependence
  [more interdependent = higher q]

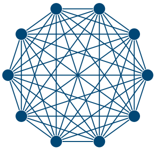
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Network Topology and IDS

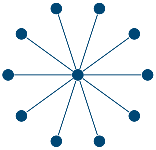
Examples



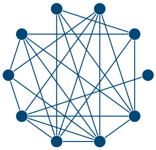
ideosyncratic



fully connected

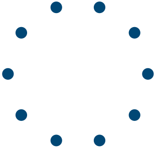


single-factor model



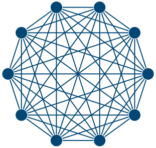
Erdös-Rényi graph

Examples



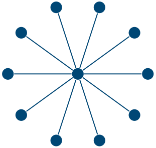
ideosyncratic

hardware failure



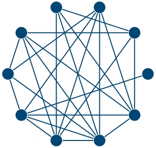
fully connected

email spam



single-factor model

OS vulnerability



Erdös-Rényi graph

inter-organizational dependence

Basic IDS model and beyond

IDS $\Longleftrightarrow$ Indirect attacks $q > 0$

Discreet security choice, N or S. (nonatomic players)

Prob. of direct breach p and of indirect q; x - fraction of insecure nodes; $q'(x) > 0, q''(x) > 0, q(0) = 0, q(1) = \bar{q} < 1$

$$B_N = p + (1-p)q(x) \quad and \quad B_S = q(x)$$

Continuous security choice (atomic and non-atomic (finite) players)

p_i - prob. of direct loss; $q_{ij} \geq 0$ - prob. of indirect attack from node j to i

$$B_i(s_i, s_{-i}) = 1 - s_i \prod_{j \neq i}^{n} (1 - (1 - s_j) q_{ij})$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Notation and Player objectives: Binary security decision

s	state $s = \{S, N\}$ (Secure, Not secure)
p	prob. of direct loss
q	prob. of indirect loss $q'(x) > 0, q''(x) > 0, q(0) = 0, q(1) = \bar{q} < 1$
W	initial wealth
L	size of a loss
$U(w)$	agent's utility with wealth w; $U'(\cdot) > 0$; $U''(\cdot) < 0$
c_i	player i cost of self-protection for $s = S$ ($p = 0$)

$$V(x, c_i) = \max_{s=\{S,N\}} p\left[1 - I_s\right] \underline{U} + \left(1 - p\left[1 - I_s\right]\right) \times \left\{q(x)\underline{U} + (1 - q(x))\bar{U}\right\} - c_i I_s,$$

$$\bar{U} := U(W); \quad \underline{U} := U(W - L); \quad I_s = \begin{cases} 1 & \text{if s=S} \\ 0 & \text{if s=N} \end{cases}$$

With no self protection

$$V(x) = p\underline{U} + (1 - p)\left\{q(x)\underline{U} + (1 - q(x))\bar{U}\right\} \quad if \quad s = N$$

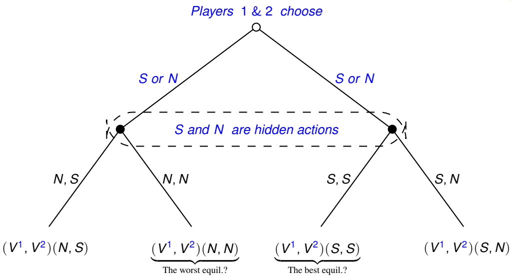
With self-protection $(p = 0)$:

$$R(x, c_i) = q(x)\underline{U} + (1 - q(x))\bar{U} - c_i \quad if \quad s = S$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Game 1: Secure or not?

Both nodes (players) simultaneously decide to *secure*(S) or *not*(N)

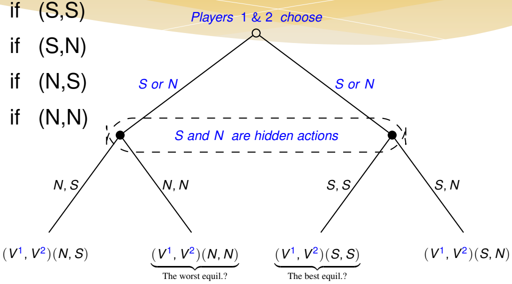


$$V^i = B_s \underline{U} + (1 - B_s)\bar{U} - cI_s, \quad where \;\; B_s = \begin{cases} p + (1-p)q(x) & if \;\; s = N \\ q(x) & if \;\; s = S \end{cases} \quad and \;\; I_s = \begin{cases} 0 & \text{if s=N} \\ 1 & \text{if s=S} \end{cases}$$

Game 1: Two identical players: secure or not?

$$V^1 = \begin{cases} \bar{U} - c & \text{if } (S,S) \\ \bar{q}\underline{U} + (1-\bar{q})\bar{U} - c & \text{if } (S,N) \\ p\underline{U} + (1-p)\bar{U} & \text{if } (N,S) \\ p\underline{U} + (1-p)\left\{\bar{q}\underline{U} + (1-\bar{q})\bar{U}\right\} & \text{if } (N,N) \end{cases}$$

$$q(0) = 0, q(1) = q(2) = \bar{q}$$



Theorem

There exists an equilibrium of the game. At most, there exists 2 equilibria: (S,S) and (N,N). Then:

- *If a player believes that another player is secure, he will secure.* $\Longleftrightarrow$ *equil.* (S, S)
- *If a player believes that another player is insecure, he will not secure* $\Longleftrightarrow$ *equil.* (N, N)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Game 1: Equilibrium

1 \ 2	S	N
S	$\bar{U}-c$,(same)	$\bar{q}\underline{U}+(1-\bar{q})\bar{U}-c, p\underline{U}+(1-p)\bar{U}$
N	$p\underline{U}+(1-p)\bar{U}, \bar{q}\underline{U}+(1-\bar{q})\bar{U}-c$	$p\underline{U}+(1-p)\{\bar{q}\underline{U}+(1-\bar{q})\bar{U}\}$, (same)

1 \ 2	S	N
S	$\bar{U}-c$, (same)	$\bar{U}-\bar{q}\Delta U-c, \bar{U}-p\Delta U$
N	$\bar{U}-p\Delta U, \bar{U}-\bar{q}\Delta U-c$	$\bar{U}-(p+\bar{q})\Delta U$, (same)

$$q(0)=0, q(1)=q(2)=\bar{q}, \Delta U=\bar{U}-\underline{U}$$

Theorem

There exists an equilibrium of the game. At most, there exists 2 equilibria: (S,S) and (N,N). Then:

- *If a player believes that another player is secure, he will secure.* $\Longleftrightarrow$ *equil.* (S,S)
- *If a player believes that another player is insecure, he will not secure* $\Longleftrightarrow$ *equil.* (N,N)

Game 2: Nonatomic identical players: secure or not?

$$V = \begin{cases} q(x)\underline{U} + (1 - q(x))\bar{U} - c & \text{if} \quad \text{(S,q(x))} \\ p\underline{U} + (1-p)\left\{q(x)\underline{U} + (1 - q(x))\bar{U}\right\} & \text{if} \quad \text{(N,q(x))} \end{cases}$$

Theorem

There exists an equilibrium of nonatomic game with identical players, and it is symmetric: (S,S) or (N,N). If

$$q(x) \leq 1 - \frac{1}{[\bar{U} - \underline{U}]} \frac{c}{p},$$

everyone invests in self-protection.

Corollary

Let there exists $x^* < 1$, s.t.:

$$q(x^*) := 1 - \frac{1}{[\bar{U} - \underline{U}]} \frac{c}{p}.$$

Then, (S, S) will be socially efficient equilibrium. Let there be common knowledge that some fraction of population x_b believes that others do not invest in self-protection. If $(1 - x_b) > x^*$, then (N, N) will be an equilibrium supported by such beliefs.

Game 3: Nonatomic players with different security costs

c_i	cost of player i, $c_i \in [c_{min}, c_{max}]$
$F(c)$	distribution function of agents' costs of protection
$f(c)$	density of $F(c)$

$$V(x, c_i) = \max_{s=\{S,N\}} p\left[1 - I_s\right] \underline{U} + \left(1 - p\left[1 - I_s\right]\right) \times \left\{q(x)\underline{U} + (1 - q(x))\bar{U}\right\} - c_i I_s,$$

$$\bar{U} := U(W); \quad \underline{U} := U(W - L); \quad \Delta U := \left[\bar{U} - \underline{U}\right].$$

$$V^i = \begin{cases} q(x)\underline{U} + (1 - q(x))\bar{U} - c_j & \text{if} \quad \text{(S,q(x))} \\ p\underline{U} + (1 - p)\left\{q(x)\underline{U} + (1 - q(x))\bar{U}\right\} & \text{if} \quad \text{(N,q(x))} \end{cases}$$

Proposition

For any $q(x)$, a player with a cost c_i invests in self-protection if $c_i \leq p(1 - q(x))\Delta U$.

Theorem

Generically, in Nash equilibrium there exists c^*, such that players with $c < c^*$ invest, and with $c \geq c^*$ - do not invest in self-protection. Socially optimal cut-off c^{so} for investing in self-protection is strictly higher than the individually optimal one: $c^{so} > c^*$.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

IDS with continuous actions: capturing the tradeoffs

Modeling defender incentives

- Costs of security $h(s)$
  - monetary [non-separable utility]
  - time or effort [separable utility]
- Benefits of security
  - reduced prob. of a breach $B_i = B_i(\mathbf{s})$
  - reduced size of a loss L_i

Modeling attacker incentives

- Costs of attacking [c_j]
  - monetary (equipment)
  - know-how (skills)
  - time and/or effort
- Costs of being caught
  - prob. of punishment [μ]
  - severity of punishment [$U(w_0) = 0$]
- Benefits of attacking [G_j]
  - pecuniary
  - savings (time, effort)
  - mental
    (ex. ideology, social cohesion)
    ex. Watch-Dogs game $\Longrightarrow$
    increased interest in hacking of real DOT systems
    https://games.slashdot.org/story/14/06/07/2052241/
    report-watch-dogs-game-may-have-influenced-highway-sign-ha

Player choices: (i) their types (attacker or defender) and (ii) amount of investment in security (determines sec. level)

Continuous security decisions: Notation and Objectives

n	number of players	c_j	player j cost of attack
s_i	player i security	p	prob. of direct loss
$\mathbf{s}$	state $\mathbf{s} = (s_1, ..., s_n)$	q_{ij}	i indirect loss propagated from j
W	initial value (wealth)	μ	prob. of capture of malicious user
L	size of loss	$U(w_0)$	utility if punished $U(w_0) = 0$
$h(s)$	security cost function		
$U(w)$	utility of wealth w		

Defender objective [to maximize his expected utility V_i]

$$V_i = \underline{U} + (1 - B_i) \cdot \Delta U - h(s_i)$$

$$B_i(s_i, s_{-i}) = 1 - s_i \prod_{j \neq i}^{n} (1 - (1 - s_j) q_{ij}); \quad \bar{U} := U(W); \quad \underline{U} := U(W - L); \quad \Delta U := \left[\bar{U} - \underline{U}\right].$$

Risk averse players [standard]: $U'(\cdot) > 0; U''(\cdot) < 0$

Security cost function [standard]: $h'(\cdot) > 0; h''(\cdot) > 0; h(0) = h'(0) = 0, h(1) = \infty.$

Attacker objective [to maximize his expected utility V_j]

$$V_j = (1 - \mu) U(G_j) + \mu U(w_0) - h(s_j) - c_j, \quad G_j(M, \mathbf{s}) = \frac{\sum_{i \neq j} B_i(\mathbf{s}) L}{M}.$$

IDS as attack technology

When attack propagation is identical across links: $q_{ij} = q$

$$B_i(s_i, s_{-i}) = 1 - s_i \prod_{j \neq i}^{n} (1 - (1 - s_j)q)$$

Let $q(n)n$ remains small as n increases: $g_\infty := q(n)n|_{n \to \infty}$ – small. Ignoring the terms non-linear in q:

$$B_i = 1 - s_i + s_i q(n) \sum_{j \neq i}^{n} (1 - s_j), \tag{1}$$

Or

$$B_i = 1 - s_i + s_i q_n \left\{ (1 - \bar{s}) - \frac{(1 - s_i)}{n} \right\}, \tag{2}$$

where $q_n := q(n)n$ and

$$\bar{s} = \frac{1}{n} \sum_{j=1}^{n} s_j$$

IDS in large networks

In the limit $n \to \infty$ $\quad q_\infty := q(n)n|_{n\to\infty}$

$$B_i = 1 - s_i\left[1 - g_\infty + \tilde{s}\right], \quad \tilde{s} := g_\infty \bar{s}$$

s_i – player i security
$\tilde{s}$ – network security

Objective function of defenders (honest)

$$V_i = \underset{\sim}{U} + (1 - B_i)\cdot \Delta U - h(s_i)$$

Objective function of attackers (malicious)

$$V_j = (1-\mu)U(G_j(M,\mathbf{s})) + \mu U(w_0) - h(s_j) - c_j, \quad G_j(M,\mathbf{s}) = \frac{\sum_{i\neq j} B_i(\mathbf{s})L}{M}.$$

Definition (Nash Equilibrium of the game Γ)

A strategy profile $(M,\mathbf{s})$ is an equilibrium if there exists no unilateral payoff-improving deviation for any player of any type.

The game $\Gamma(M)$ with a fixed number of attackers

Definition (Nash Equilibrium of $\Gamma(M)$)

Consider the game $\Gamma(M)$ with a fixed number of attackers. A strategy profile $\mathbf{s} = (s_1, \ldots, s_N)$ is a Nash equilibrium if for every i, s_i is a best response.

Lemma

In any equilibrium of the game $\Gamma(M)$, for each user type, security choices are identical.

Theorem (Unique eq. security levels for each type)

For a given M and $h''' \geq 0$, for each player type equilibrium security $s_i^*(M)$ is unique. It is zero for attackers, and positive for defenders.

Theorem

Defender equilibrium security level decreases in the number of attackers.

FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

Equilibrium of the game Γ

Theorem

The game Γ admits at least one pure strategy Nash equilibrium.

Theorem

Socially optimal security levels s^{so} are strictly higher than the individually optimal security choices in the game Γ: $s^{so} > s^*$.

$\Longleftrightarrow$ Need to design of policies to improve security incentives.

IDS framework for large scale CPS:

- quantification of policy impact
- comparison across different policies.

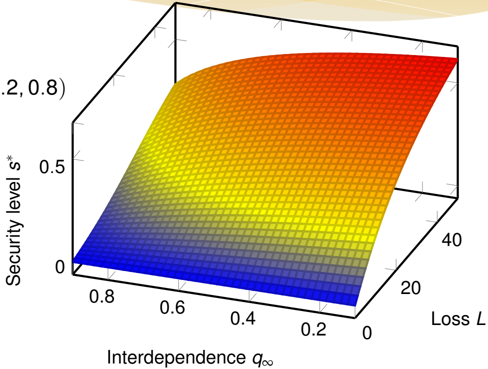
Results: Equilibrium security level

Set of parameters

$N = 500$	number of nodes
$W = 100$	initial node value
L	loss size, $L \in (0, 50)$
q	interdependence, $q \in (0.2, 0.8)$
$\mu = 0.2$	prob. of attacker capture
$U(w_0) = 0$	punishment utility

$$h(s) = 10 \frac{s^2}{\sqrt{1-s}}$$

$$U(x) = x^{0.9}$$



Eq. security level s^* as a function of L and q_∞

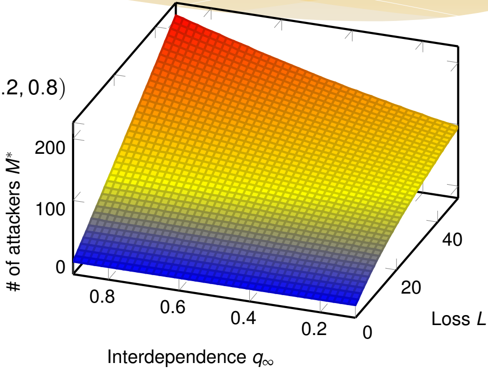
Results: Equilibrium number of attackers

Set of parameters

$N = 500$	number of nodes
$W = 100$	initial node value
L	loss size, $L \in (0, 50)$
q	interdependence, $q \in (0.2, 0.8)$
$\mu = 0.2$	prob. of attacker capture
$U(w_0) = 0$	utility in punished

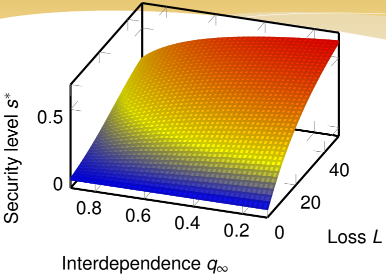
$$h(s) = 10 \frac{s^2}{\sqrt{1-s}}$$

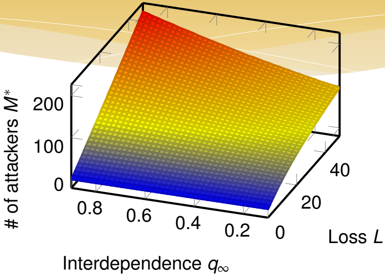
$$U(x) = x^{0.9}$$



Eq. # of attackers M^* as a function of L and q_∞.

Results: Discussion





Technology policies

- To promote technologies reducing q_∞ ?
- To mandate min security level $\hat{s}$ (required best practices?)

Policies require quantification

[of social costs and benefits based on aggregation of individual risks]

IDS framework for large scale CPS provides

- Parameter-based valuation of risks for large scale CPS systems
- Allows to consider strategic defenders and attackers

Conclusion and directions

IDS framework for large scale CPS

- Internet of things requires new tools for risk evaluation & management
- Our IDS framework
  - Evaluates risks for systems with various topologies
  - Allows to design cyber-insurance and assess its effects

Cautious optimism Global risks 2015, Global risks 2016

Technological Risks 2014 → 2015

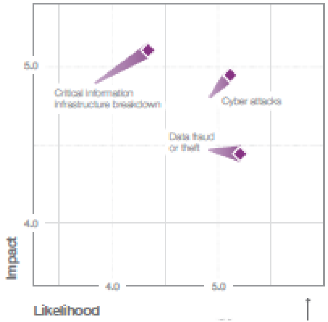


Figure 1.1: The Changing Global Risks Landscape 2015–2016: The 10 Most Changing Global Risks