Oscillations through Co-evolution: A Manifestation of Moving Target Defense Conficker Case Study

Daniel Bilar (Siege Technologies)
George Cybenko (Dartmouth College)
John Murphy (ProQueSys)

CSIIRW 8, ONRL (Oak Ridge, TN)
January 9, 2013

Support

- Research partially supported by DARPA I2O, DHS, AFRL, DOD, OSD, and AFOSR with UTEP, Ball Aerospace, Pikewerks, Siege
  - All opinions and results expressed are those of authors and not necessarily those of the funding agencies
- Thanks also to **V. Berk, I. Gregoriou-de Souza, J.T House, D. Sicilia, G. Stocco, P. Sweeney**

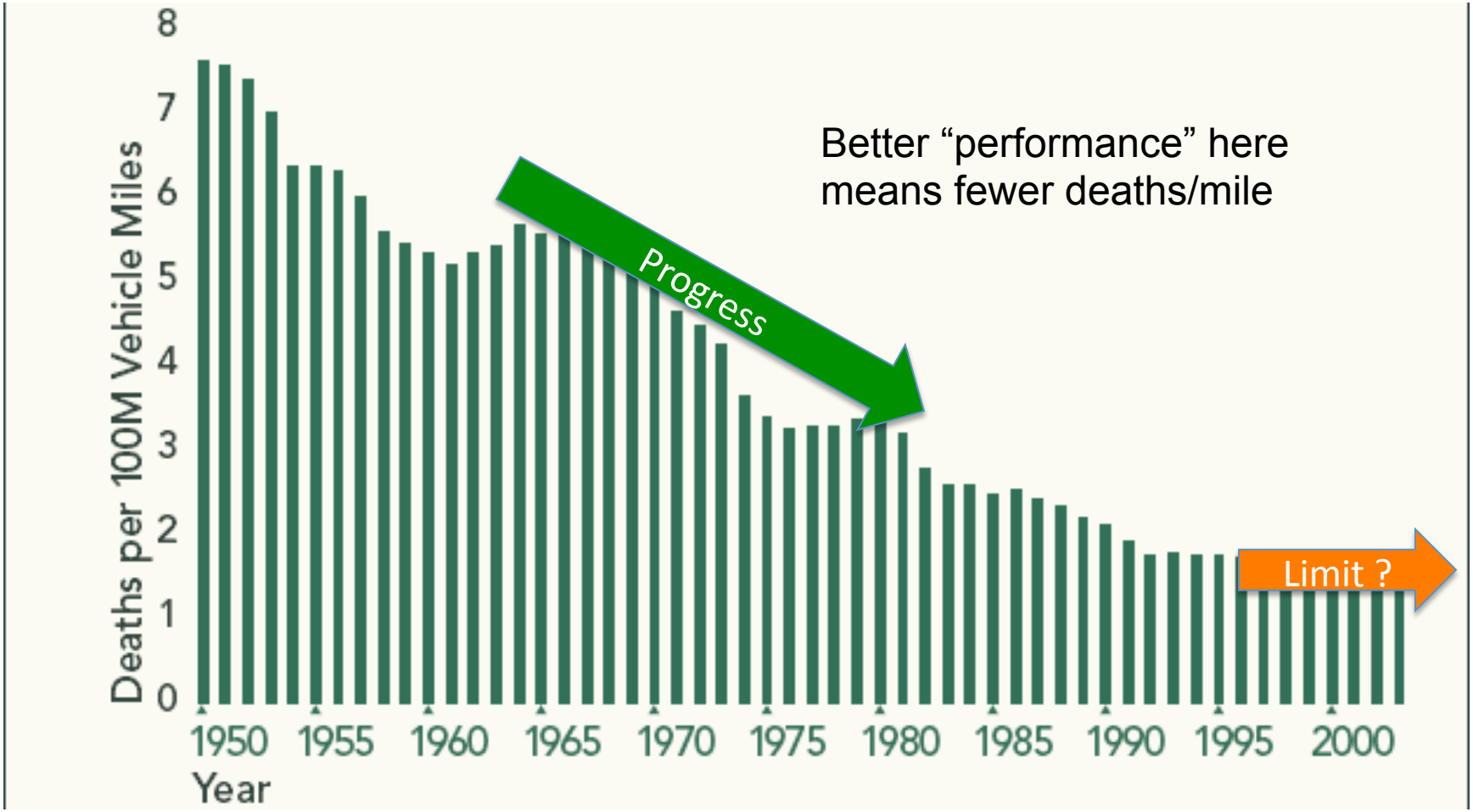
Outline of talk 1/2

- **Background**: Studied public data in various domains
  - US border security, computer vulnerability databases, offensive & defensive coevolution of worms (Conficker)
  - Modeled as players in adversarial situation
- **Findings**: Performance metrics **oscillate over time**
  - No asymptotic convergence, not monotonic
- **Claim**: In majority of (adversarial) games, players do *not* compute Nash Equilibriums over (static) strategy sets but use myopically perceived best responses at each time step
  - 'Classical' game theory is not the best fit
- **Why**: Not a *stationary* environment! Ongoing sequences of moves, countermoves, deception and strategic adaptation
  - Explains exhibited oscillations and consistent with data

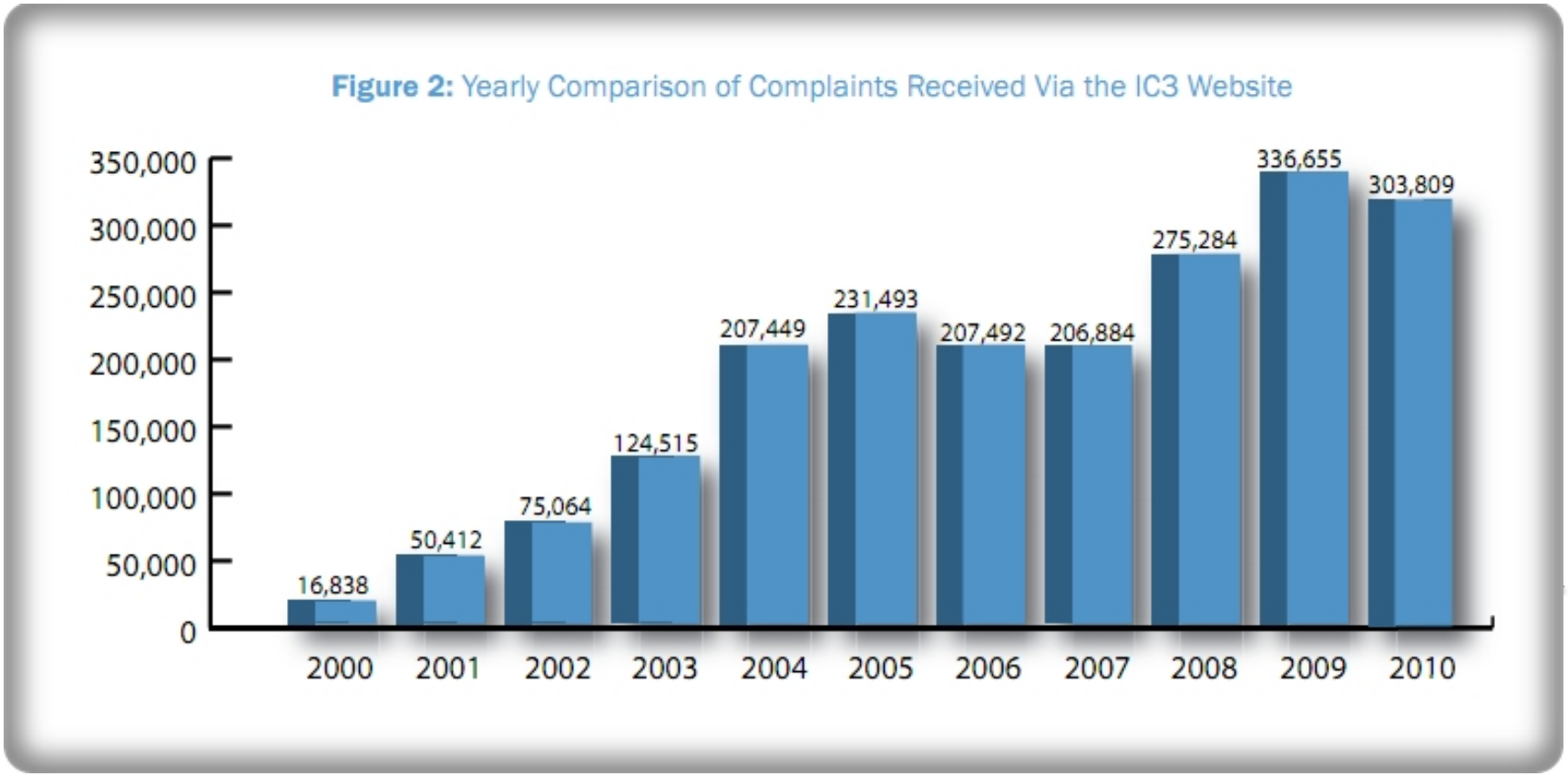
Outline of talk 2/2

- **Problem**: Oscillations modeled by replicator equations
  - Typically 3rd degree, non-linear, analytically difficult
  - Inverse problem of estimating RE parameters from observations of behavior computationally tractable
- **Claim**: Possible to infer players motives, costs and move options by observation of oscillation
  - Not discussed in this talk
- **Contributions of authors**
  - Detailed empirical analysis of players Conficker & environment (Bilar & Murphy)
  - Abstraction of game through Quantitative Attack Graph (Bilar & Cybenko & Murphy)
  - "Asymptotic" cut set theorem (Cybenko) for optimal defense allocation

You know you are working in an adversarial domain when you want to see this kind of progress…

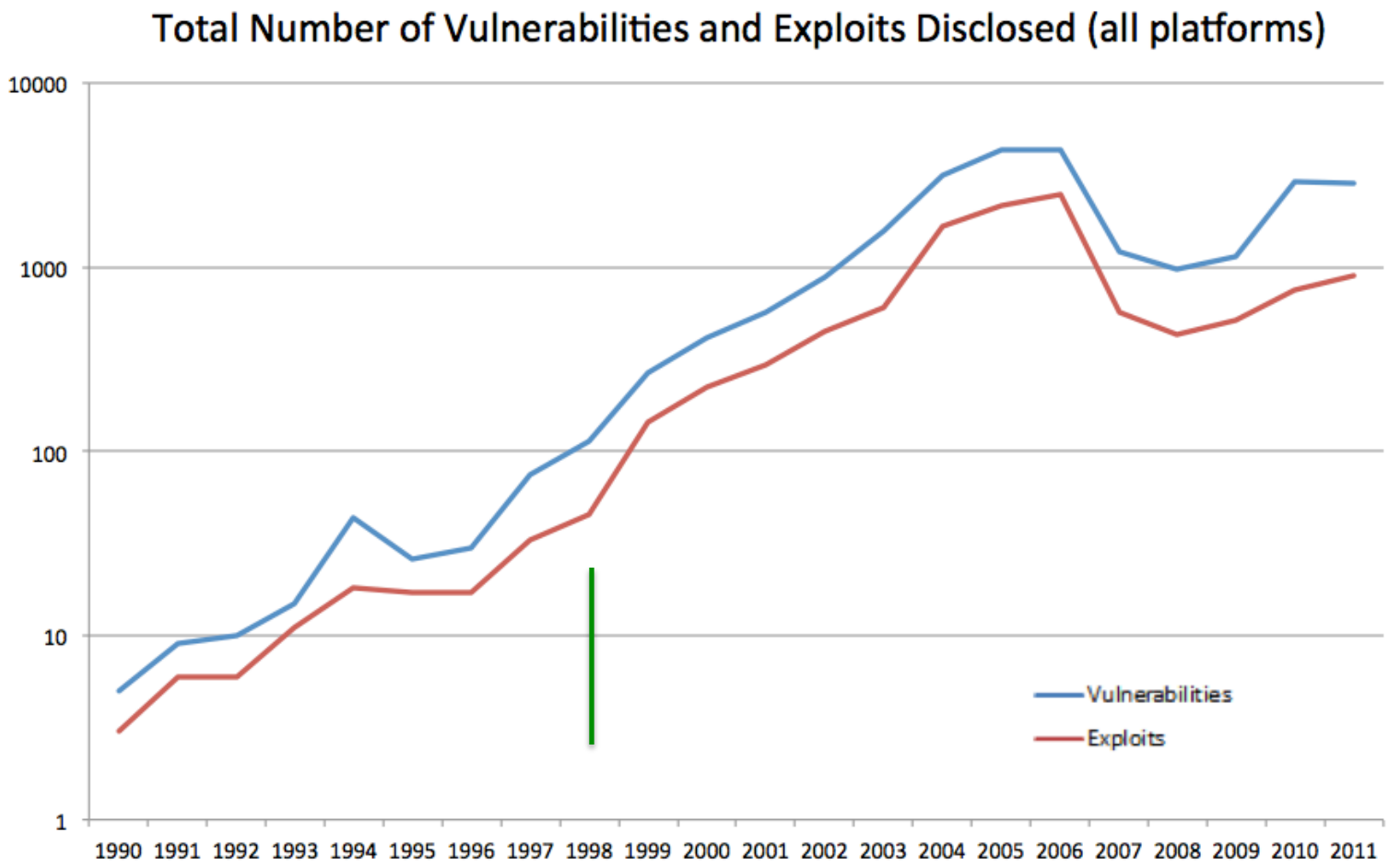


...but instead, you see this …



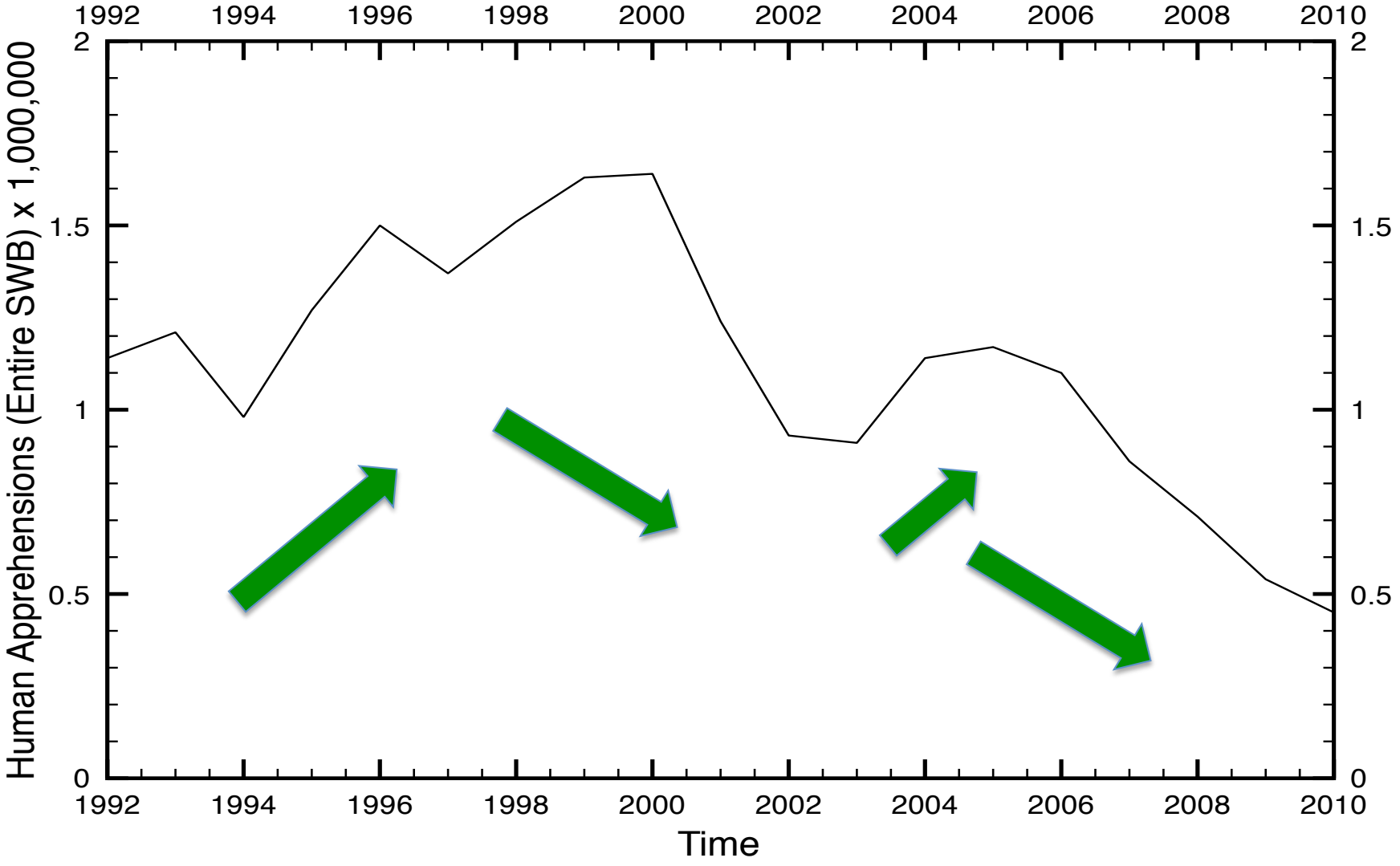
Internet Crime Complaint Center, http://www.ic3.gov/default.aspx

...or this …

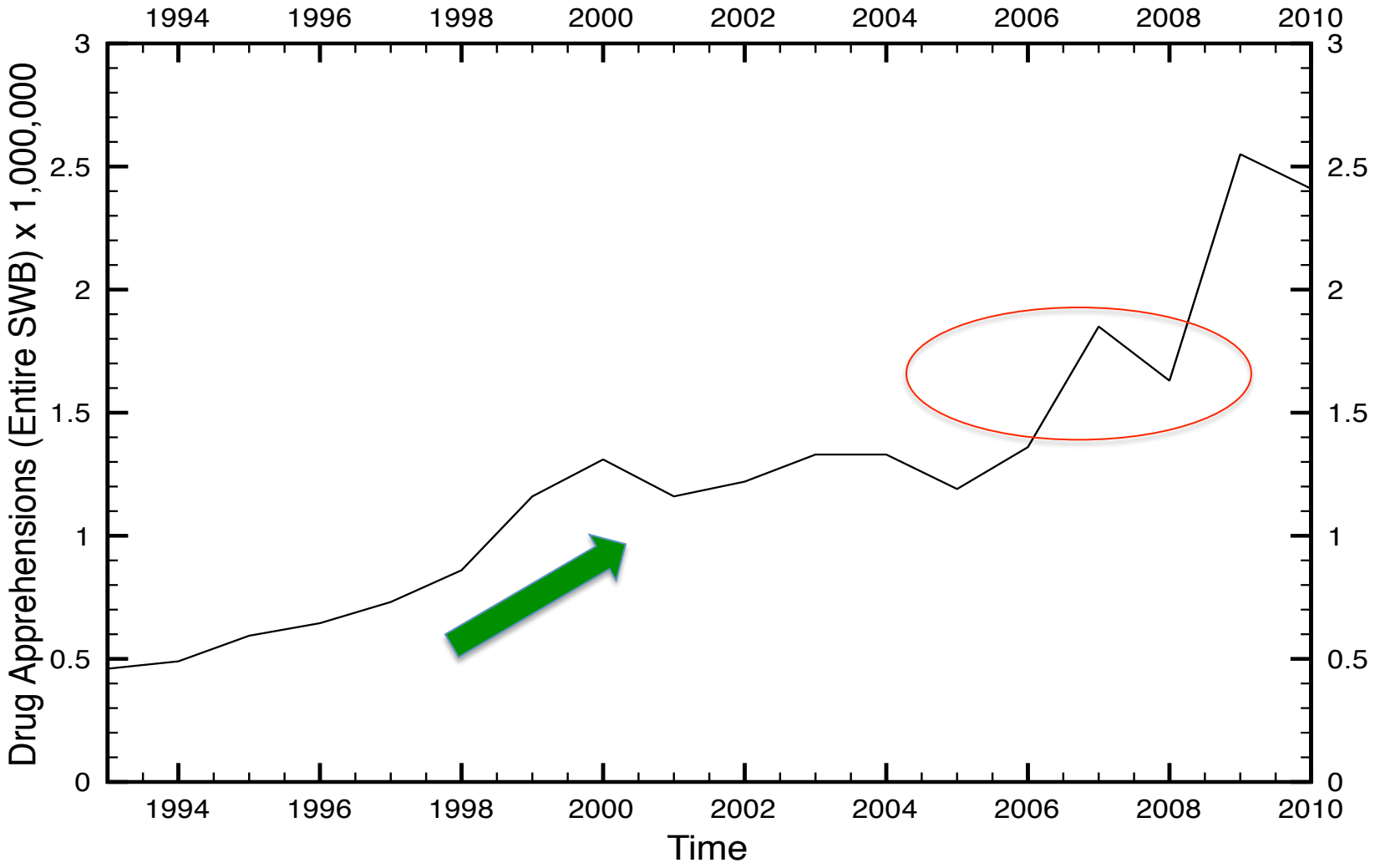


From OSVDB by P. Sweeney

Border security...



War on drugs...


1994
1996
1998
2000
2002
2004
2006
2008
2010
3
2.5
2
1.5
1
0.5
0
Drug Apprehensions (Entire SWB) x 1,000,000
Time

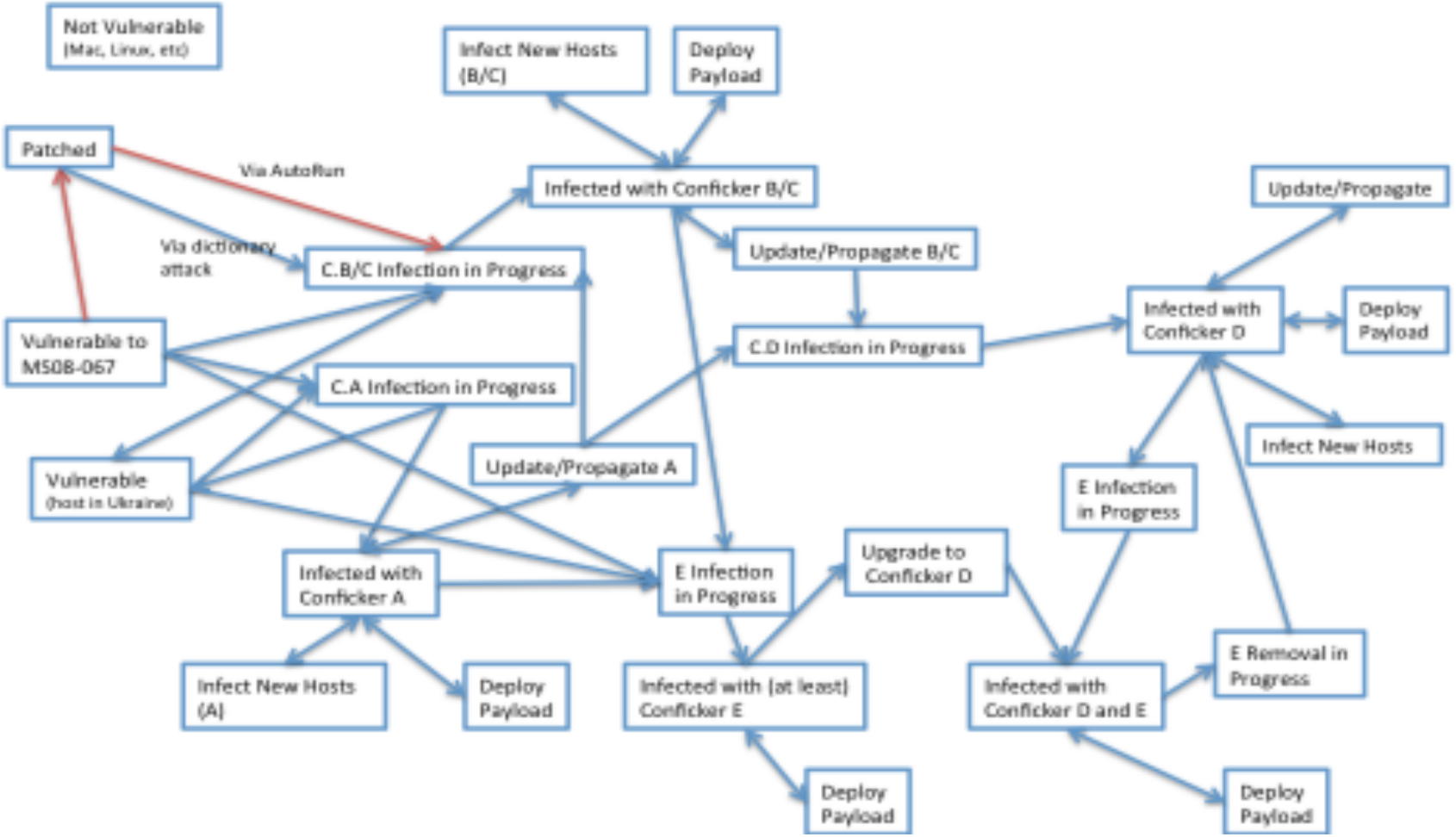
Comments

- **"Performance" measures may oscillate** (not monotonic)
  - Depends partly on normalization of metrics (see Fig 3.1 in BMC (2012))
- Operating against human adversaries is different than operating against nature
- Games *not* defined a priori, game details not known
  - Players *do not know* who the other players are, what their possible moves might be and, perhaps most importantly, what their preferred outcomes or objectives are
- Result: **Co-evolution, adaptation as evinced through oscillations**

Conficker

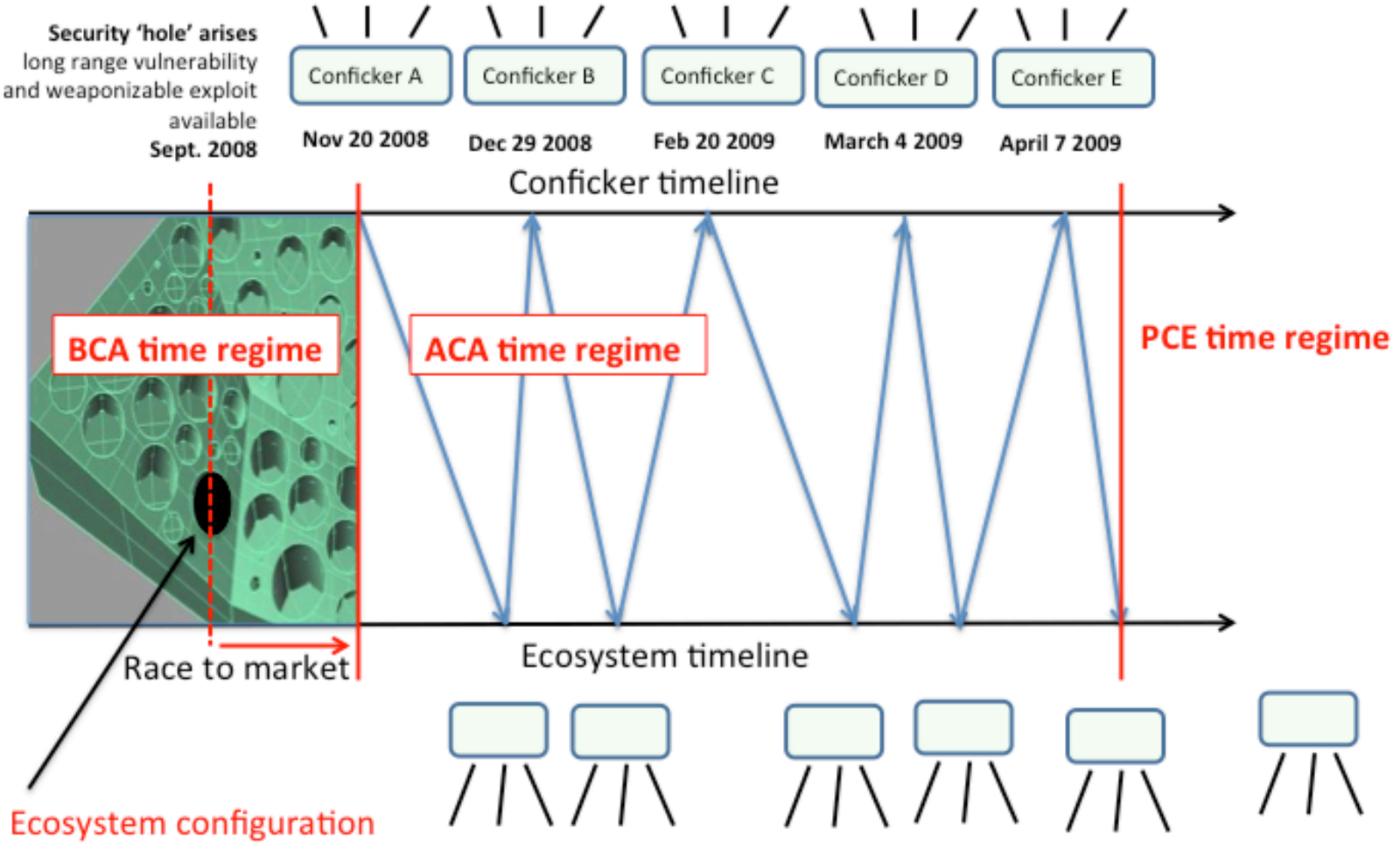
- AKA Downup, Downadup, Kido
- Detected November 2008
- Largest worm/botnet infection since 2003
- Infected million's of machines
- Evolved through 5 versions in several months
- Affected military systems in France, UK etc
- Used many vulnerabilities and techniques

Conficker Versions A-E Host States



Adversarial Dynamics: The Conficker Case Study. Daniel Bilar, George Cybenko and John Murphy
In *Moving Target Defenses II*, edited by S. Jajodia, Springer, 2012

Conficker Timeline



Examples of Conficker Analysis

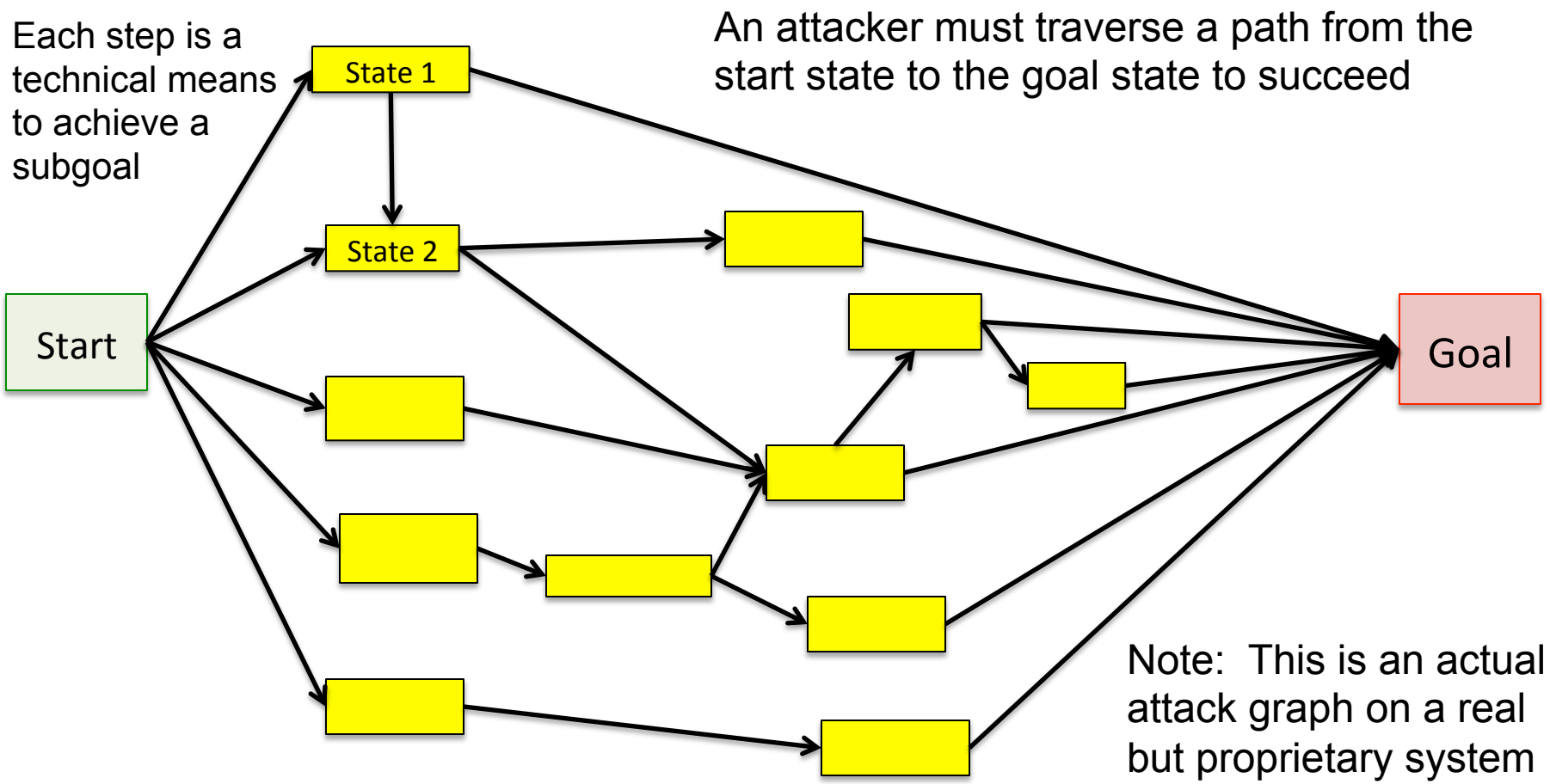
Table 3 Measures implemented by Conficker and Ecosystem between November 2008 and April/May 2009. **Bolded** indicates newly introduced measures. ~~Strike-through~~ indicates dropped measures

Time Period	Player	Spread/Infect	Update	Armor	Other
Nov 20, 2008 - Dec 28, 2008	11/20/08 Conficker.A	**MS08-067** **EnvCheck** **FetchGeoIP**	**central** **rnd250-5** **RC4** **RSA-1024**	**obfusc**	**AVXP**
	Ecosystem response	**AVsigA** **DenyGeoIP** **MSpatch**	**blockBiz**		
Dec 29, 2008 - Feb 19, 2009	12/29/08 Conficker.B	~~FetchGeoIP~~ **InclGeoIP** ~~EnvCheck~~ **LocalShare** **USB** MS08-067	~~central~~ ~~RC4~~ ~~RSA-1024~~ **rnd250-8** **MSbkcdr** **MD6v1** **RSA-4096**	obfusc **DNSblock** **AutoUpdDis** **AnlsShut**	
	Ecosystem response	**AVsigB** DenyGeoIP		**SRI-AB**	**MSBounty**
			~~MD6v1~~		

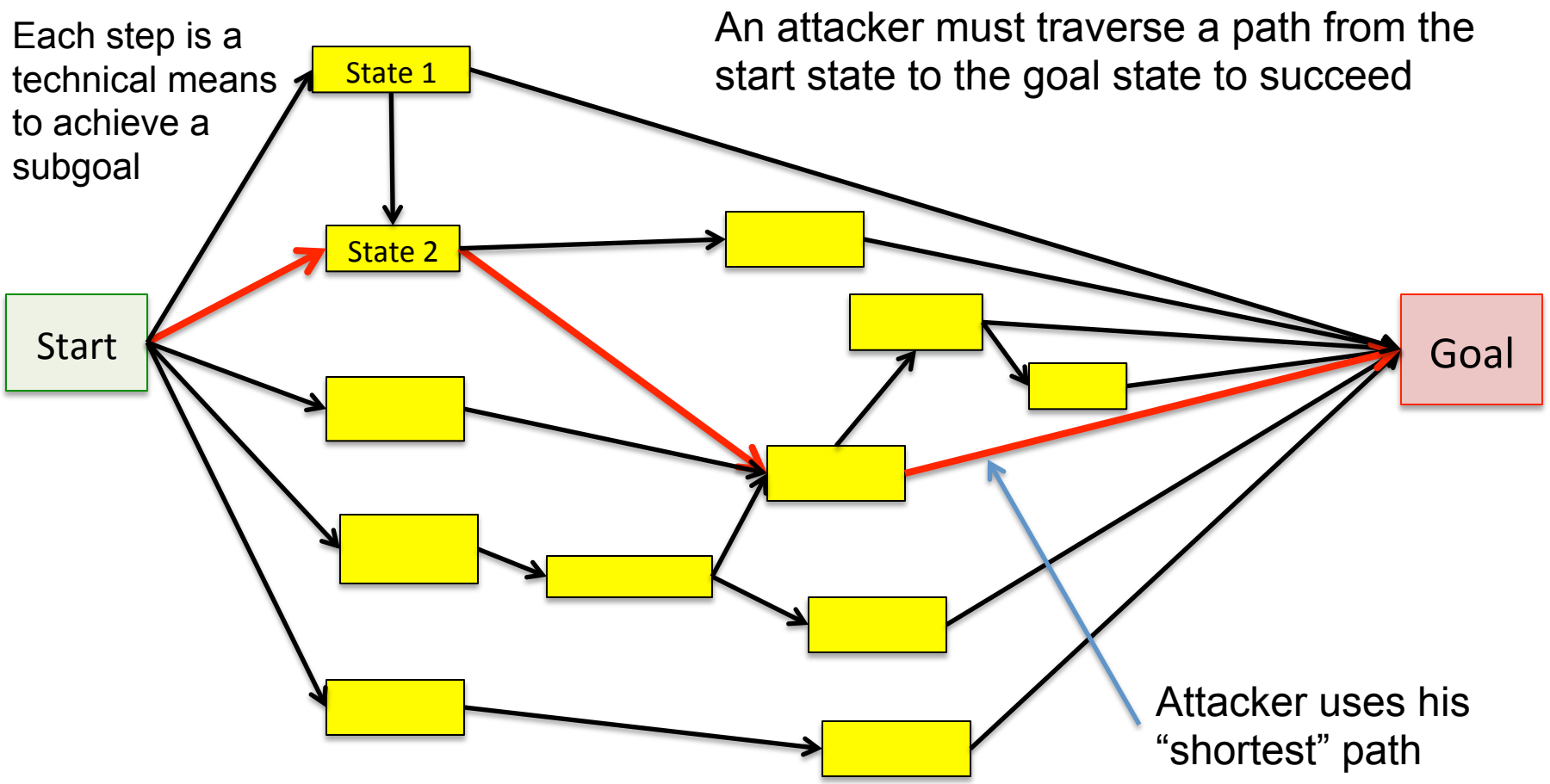
Abstraction of Attack/Defend Game

- **Attackers attacks "weakest" paths to achieve goals**
  - Weakest according to attackers' understanding
  - Paths consist of one or more technical steps
  - Can create *completely new paths* and/or steps
- Defenders make some step(s) of the most common/ damaging paths harder to traverse
  - Most common/damaging according to defenders' understanding
  - Users/boss want to create new services so new paths emerge
- Iterate the above over time

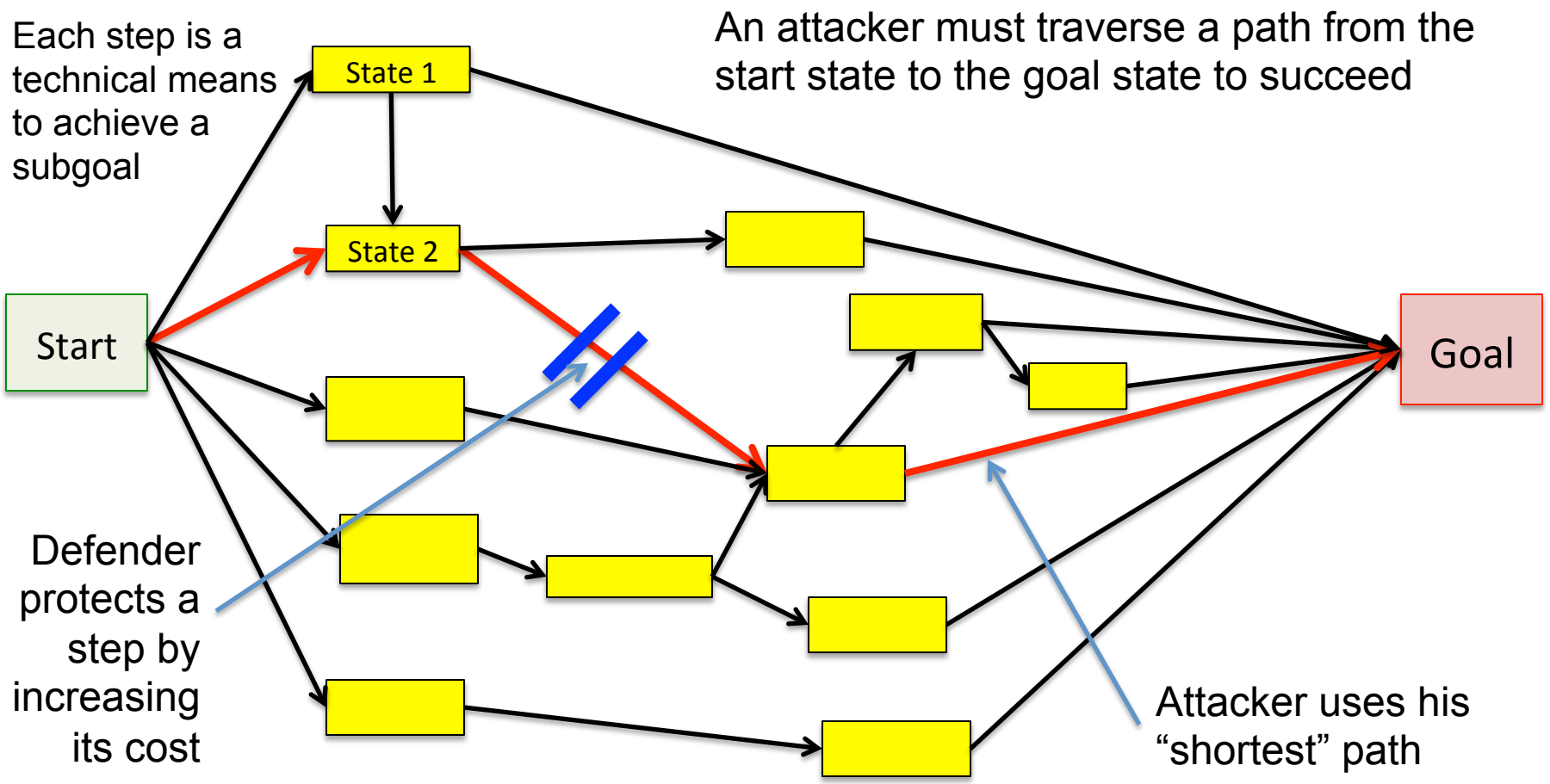
Attack Graph for a Critical System



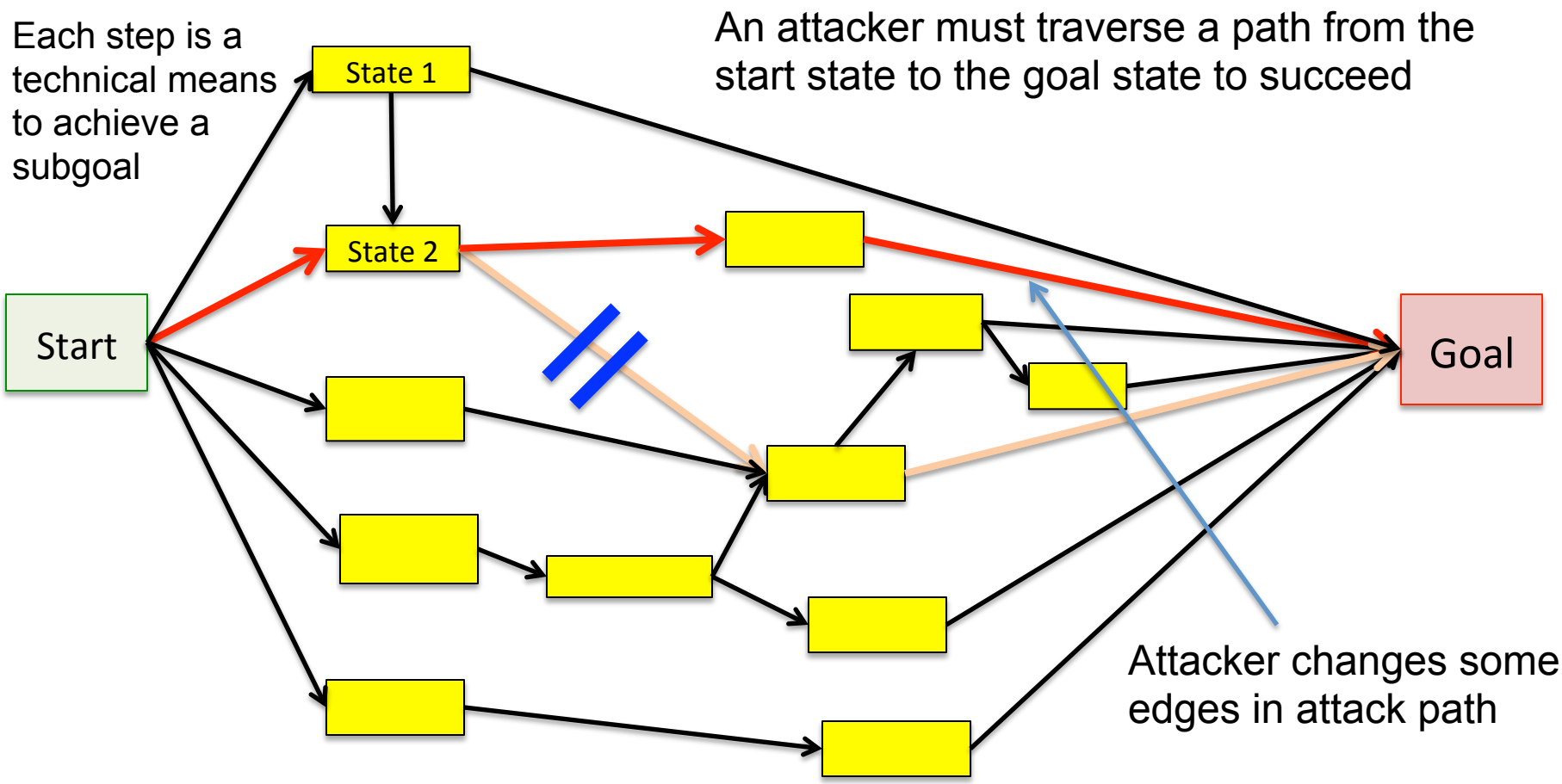
Attack Graph for a Critical System



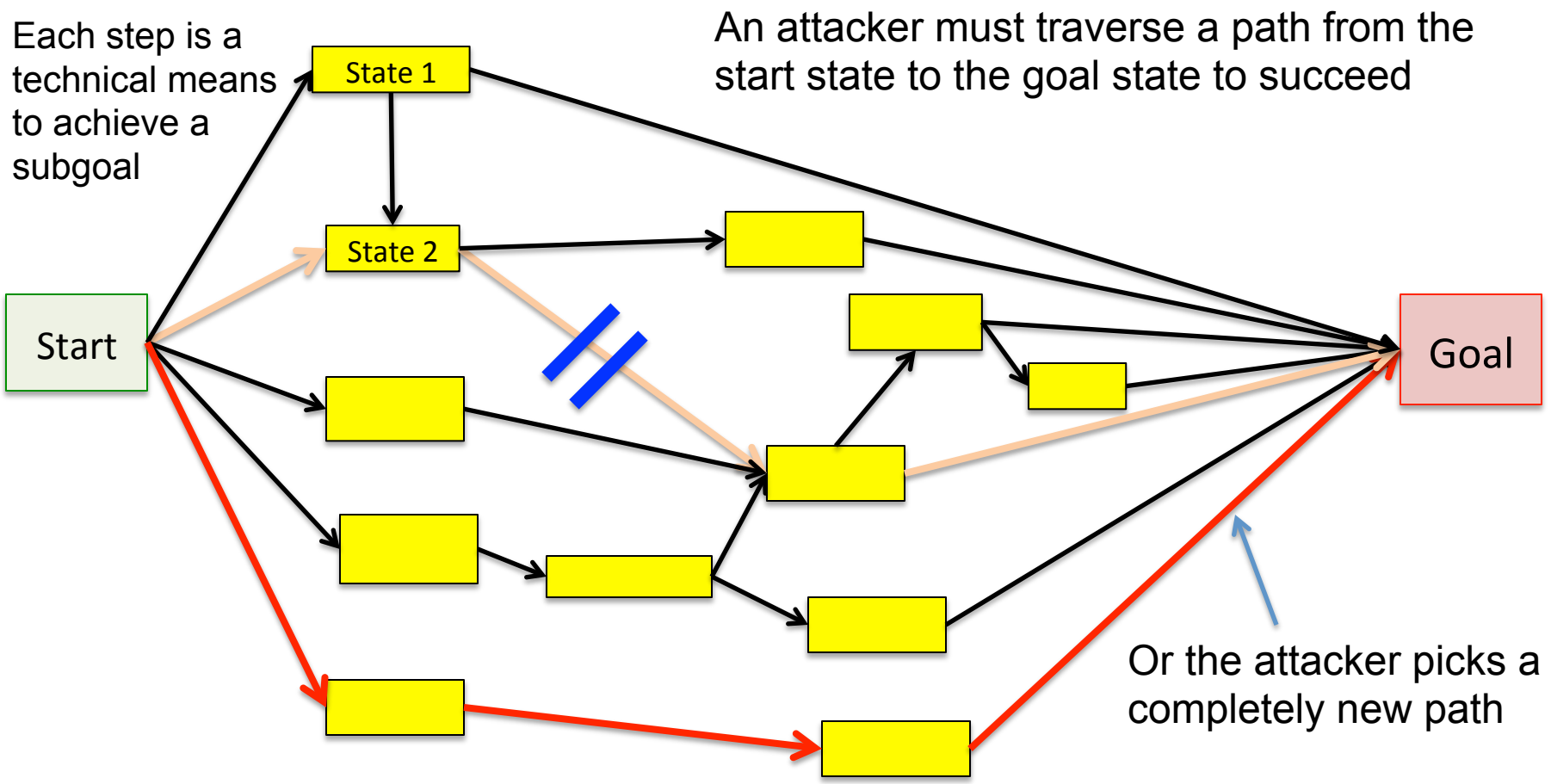
Attack Graph for a Critical System



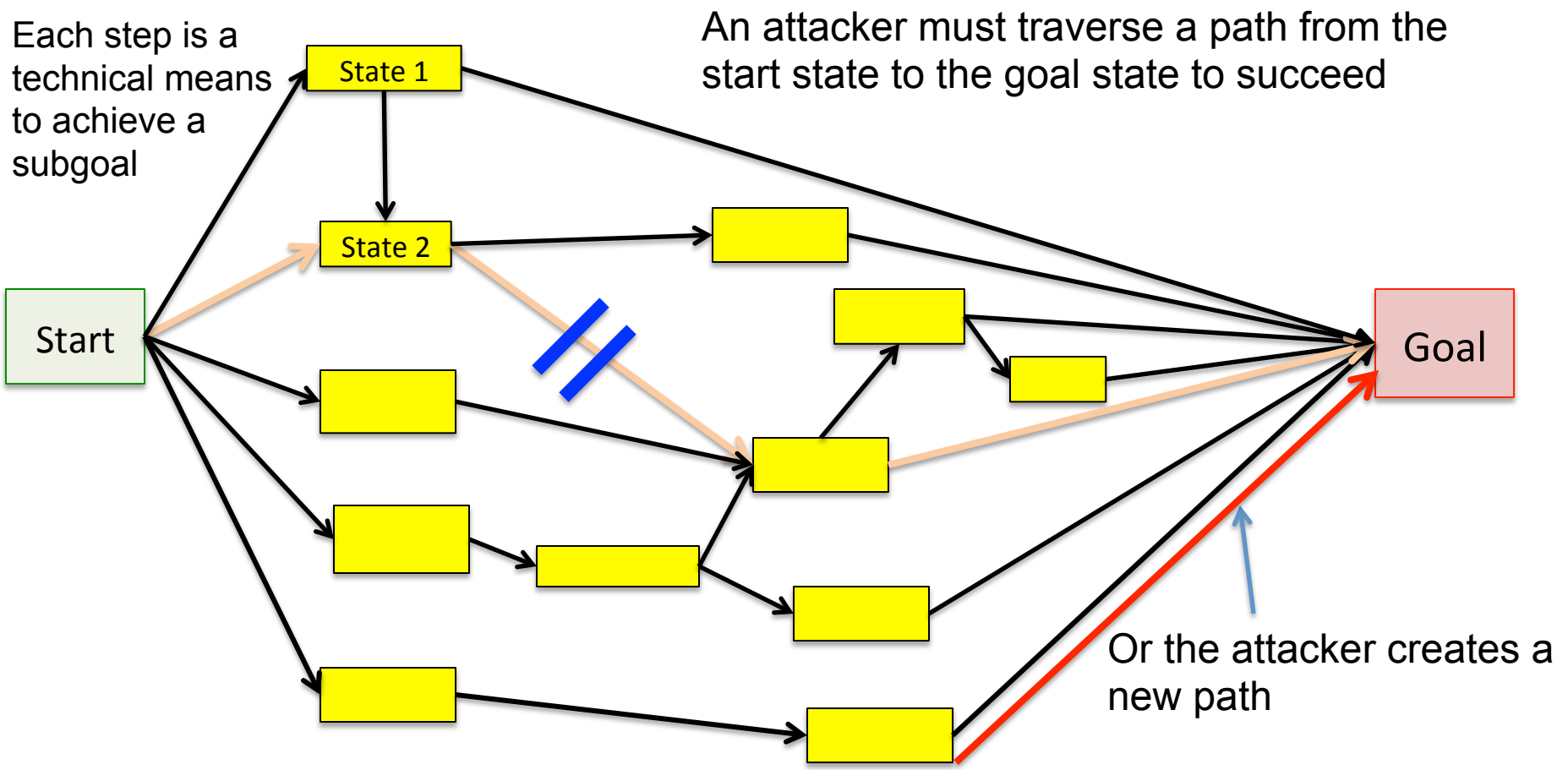
Attack Graph for a Critical System



Attack Graph for a Critical System



Attack Graph for a Critical System

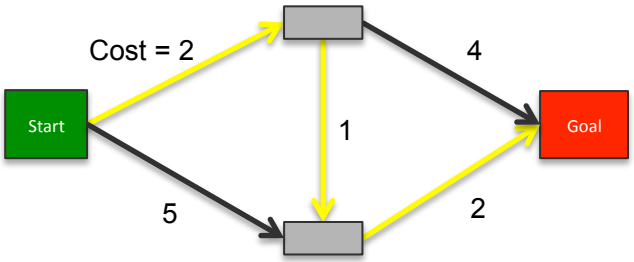


Comments

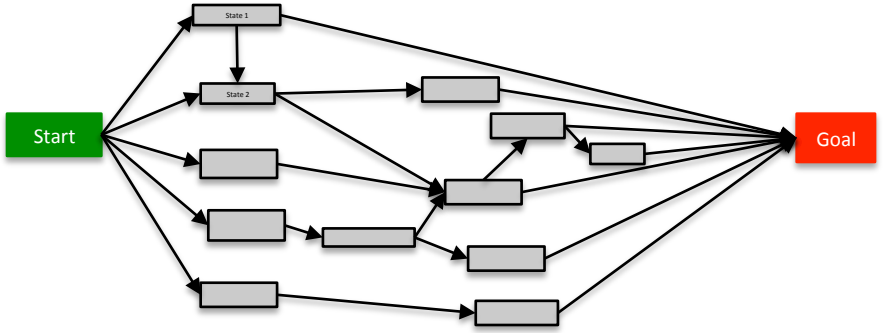
- Attacks graphs are old technique but hard to build and quantify
  - State space explosions, how to assign edge costs, blind spots, etc
  - Maybe like democracy, worst way except for all others
- Prediction markets: QuERIES provides a technique for quantifying the attack graphs by cost, difficulty, etc
- We will adapt, invest and perform better if we quantify
  - Pursuit-evasion – go to where the prey will be
  - Flu shots anticipate the flu, not respond to current ones
  - Wayne Gretzky – "A good hockey player plays where the puck is. A great hockey player plays where the puck is going to be."

Attack-Defend Game

- Estimate costs to attacker of traversing attack graph edges – shortest path is the most attractive for an attacker to take



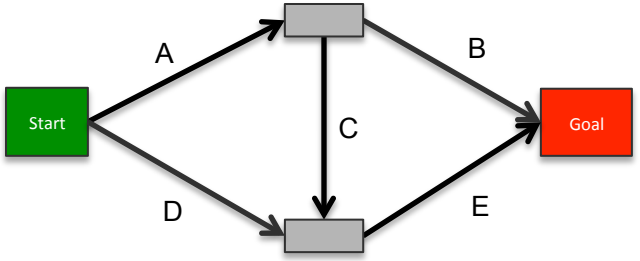
Simple Example – Shortest path in yellow



Real Problem – What is/are the shortest path(s)?

- Optimization problem – **maximize the cost of the shortest path from Start to Goal states**
- Can formulate this as a linear programming problem – solution is the **investment allocation that makes the least cost attack as expensive as possible**

Linear Programming Formulation



5 edges
3 paths

$$M = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

One column per edge
One row per path

$$u = \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix}$$

Vector of initial edge costs

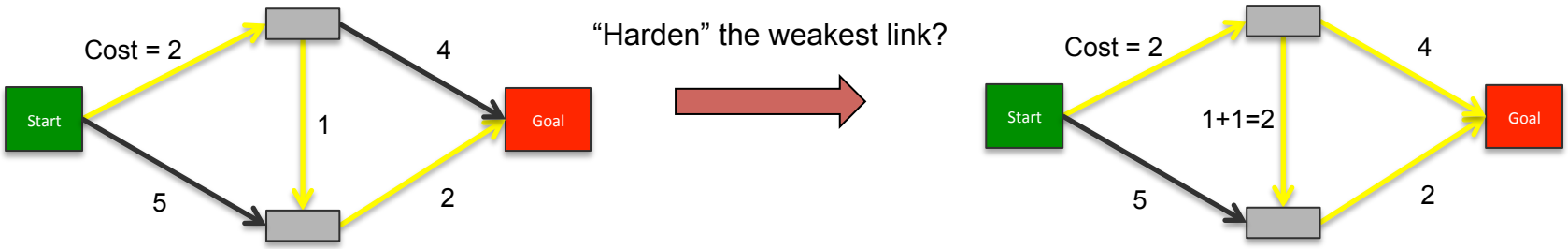
$$x = \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}$$

Vector of allocated costs

max z
such that
$M*(u+x) \geq z \geq 0$
$\mathbf{1}* x = K > 0, x \geq 0$

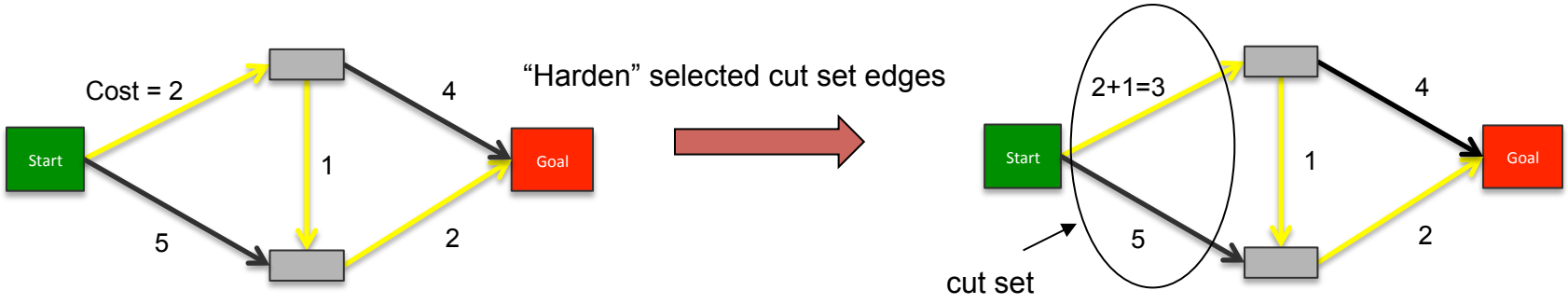
Example strategies

- Which edges are “best” to invest in? Suppose budget = 1.



Simple Example – Shortest path in yellow

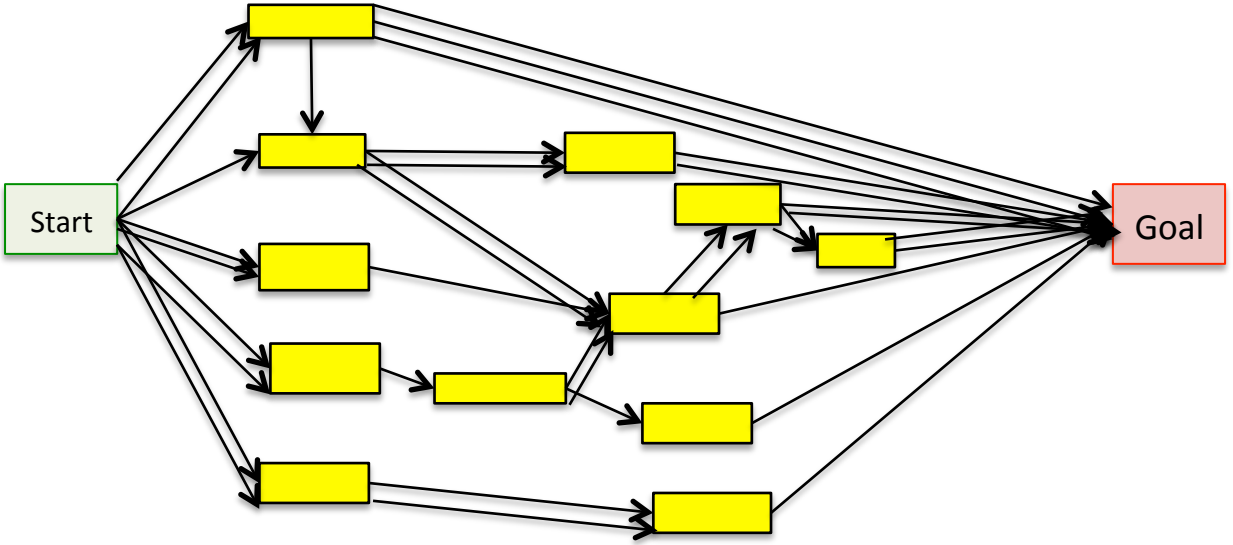
- Analysis has shown that optimal investments are <u>ultimately</u> in a “cut set”



Simple Example – Shortest path in yellow

If possible, invest in minimal cut set edges

Back to Real System

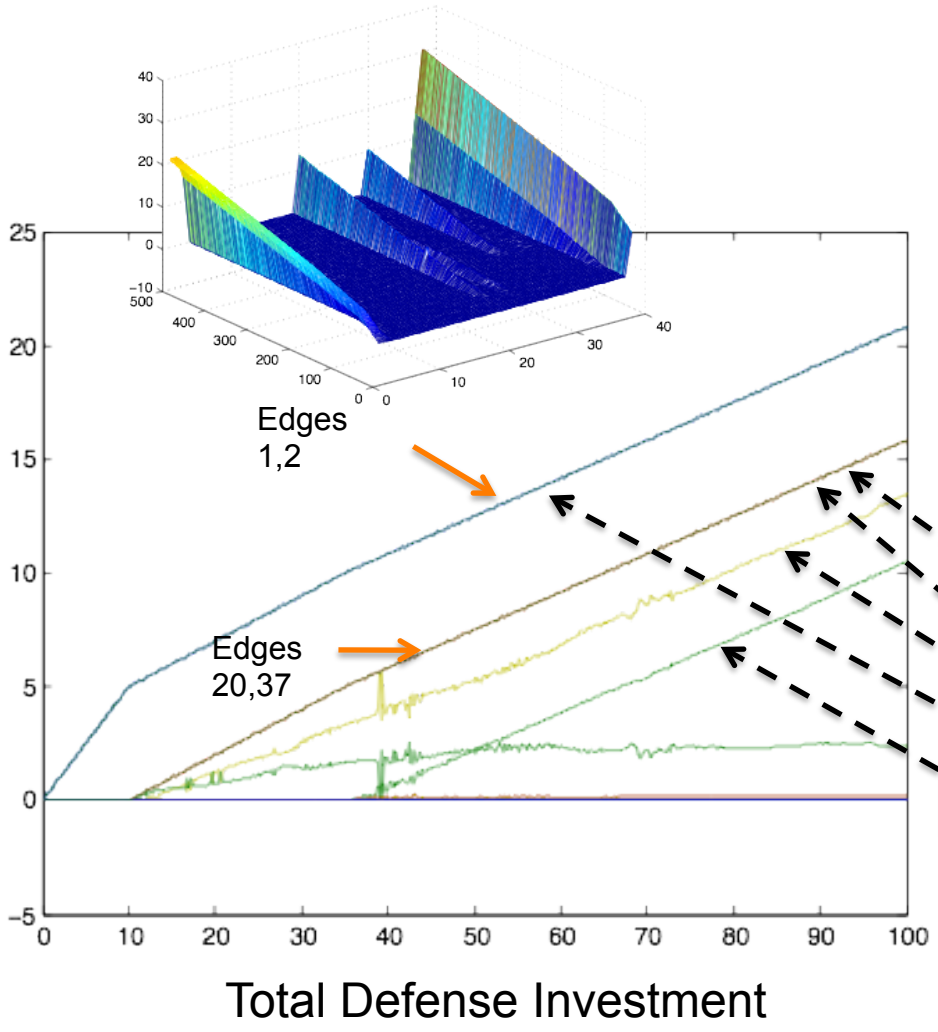


37 edges
180 paths
12 nodes

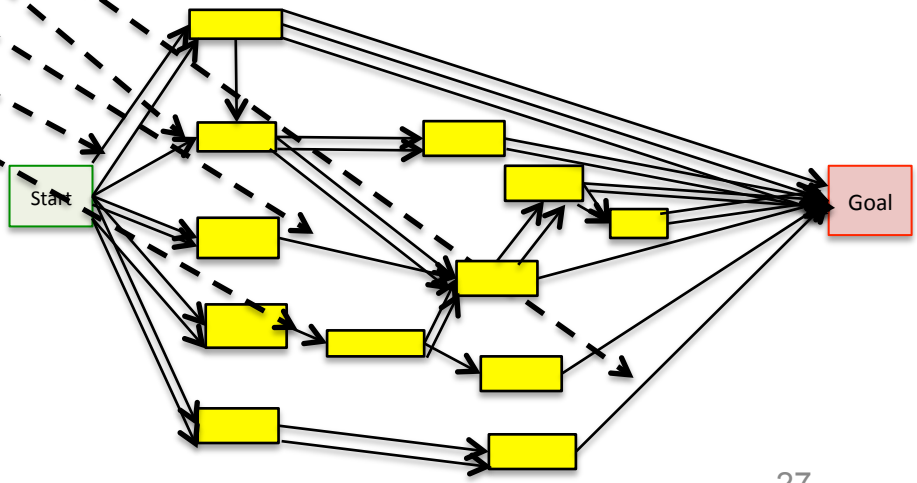
Multiple edges mean multiple attack steps possible

Matrix M has 37 columns and 180 rows

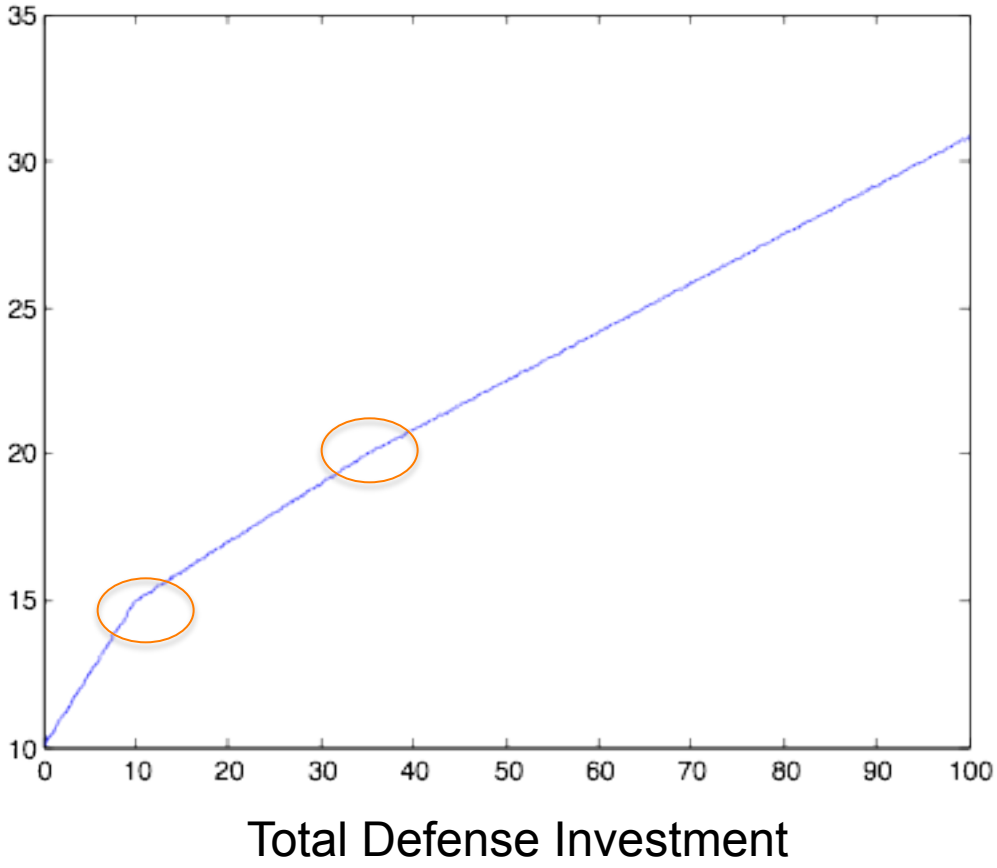
Linear Programming Results Identify High Value Protection Paths for Different Investment Levels



- Result shows benefit from hardening multiple paths according to iterative algorithm
- X-axis shows total budget, Y-axis shows investment in hardening specific paths
- As budget increases, the defensive strategy is diversified, but investment into minimal cut edges continues
- Once the inputs to state 2 are hardened, investment begins in edges 20 and 37

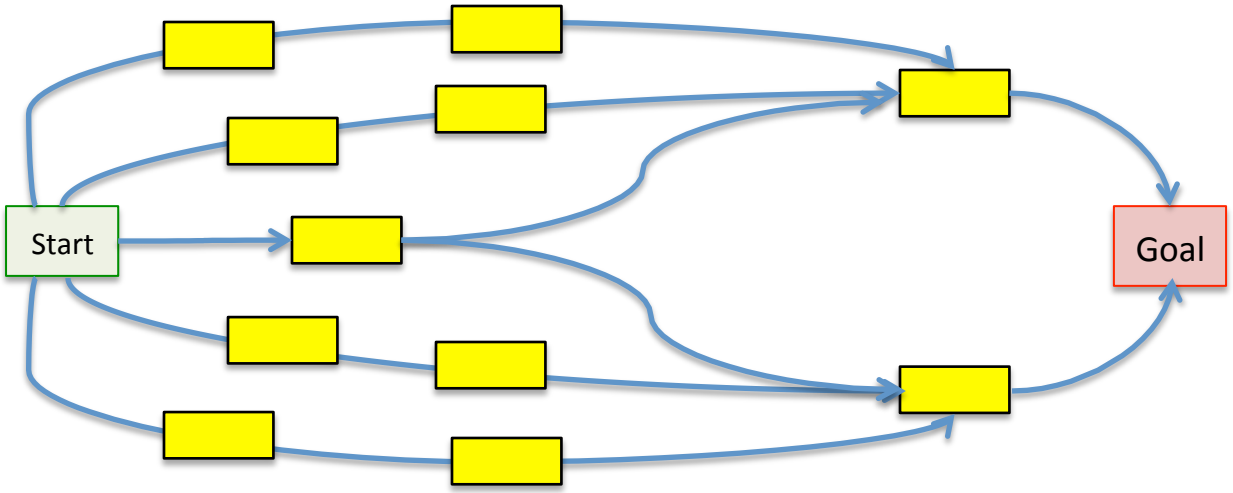


Minimal cost paths for attacker



- Graph shows total cost of minimum-cost path resulting from investment strategy
- Minimum effort required by attacker
- Includes initial edge costs along path
- Slope decreases as investment strategy diversifies into hardening multiple paths
- "Diminishing rate of return", ROI

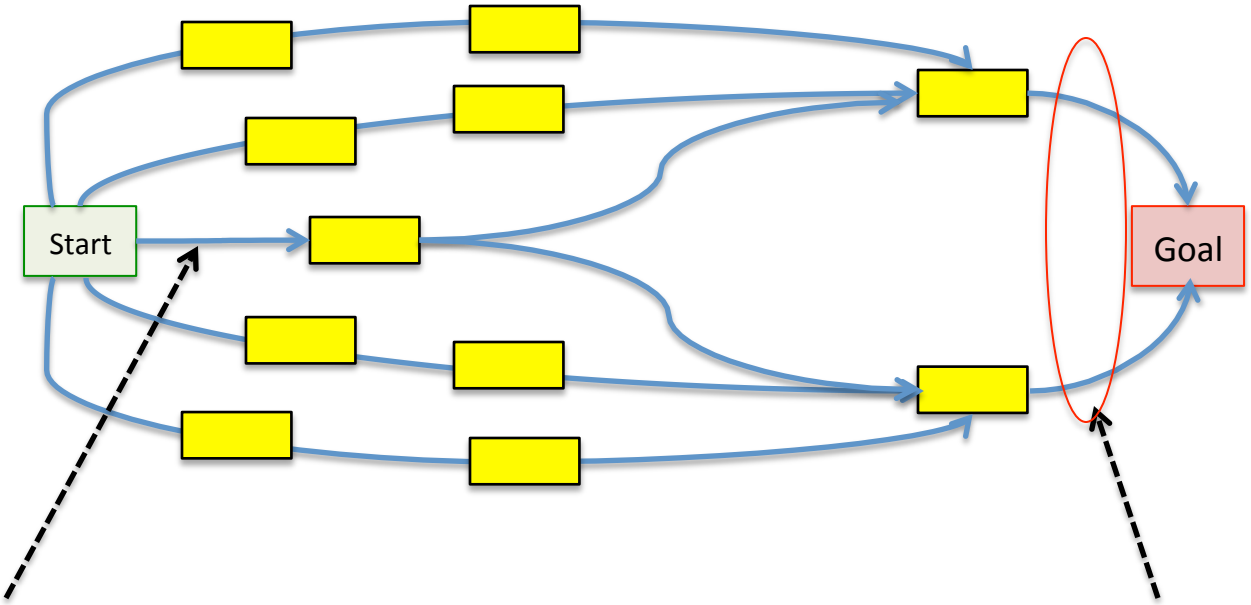
Role of minimal cut sets



Each edge has cost 1

You have a budget of 1

Role of minimal cut sets

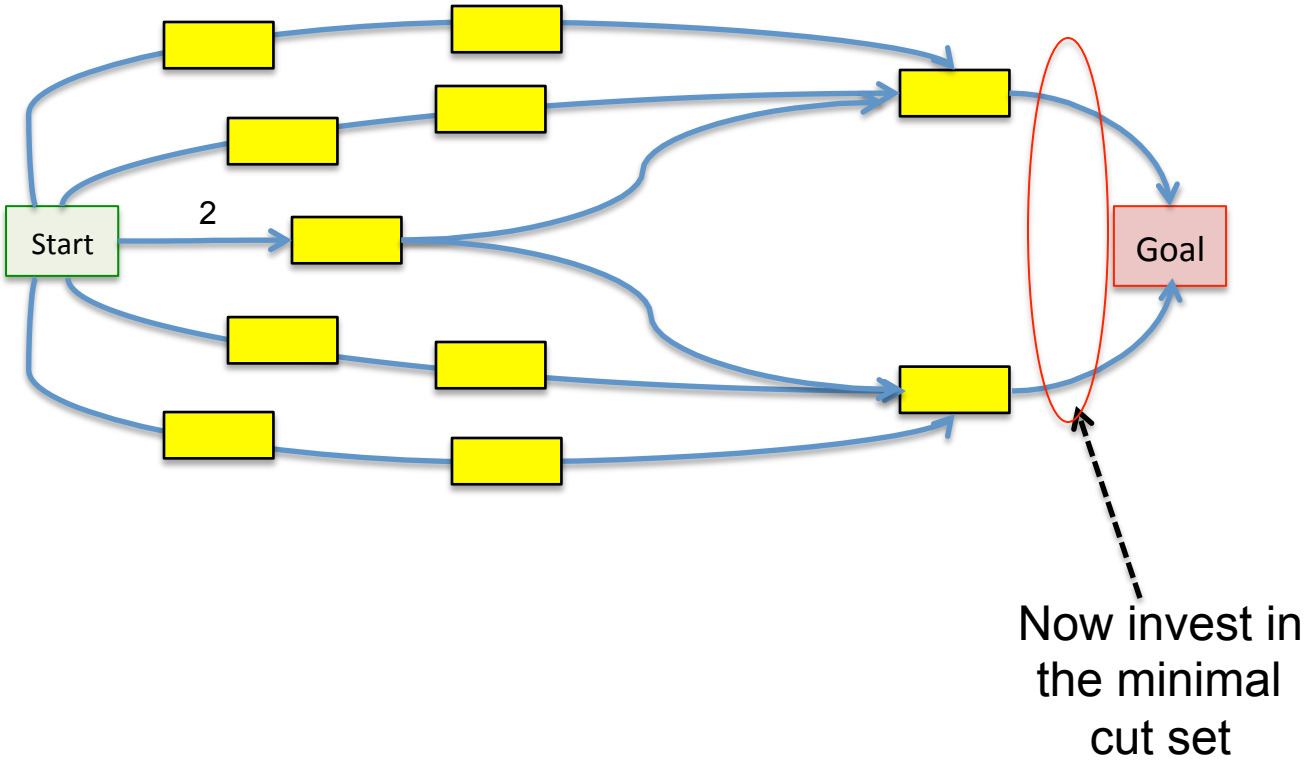


Invest that 1 unit here

But this is the minimal cut set

Each edge has cost 1
You have a budget of 1

Role of minimal cut sets



"Asymptotic" Attack Graph Theorem (Cybenko)

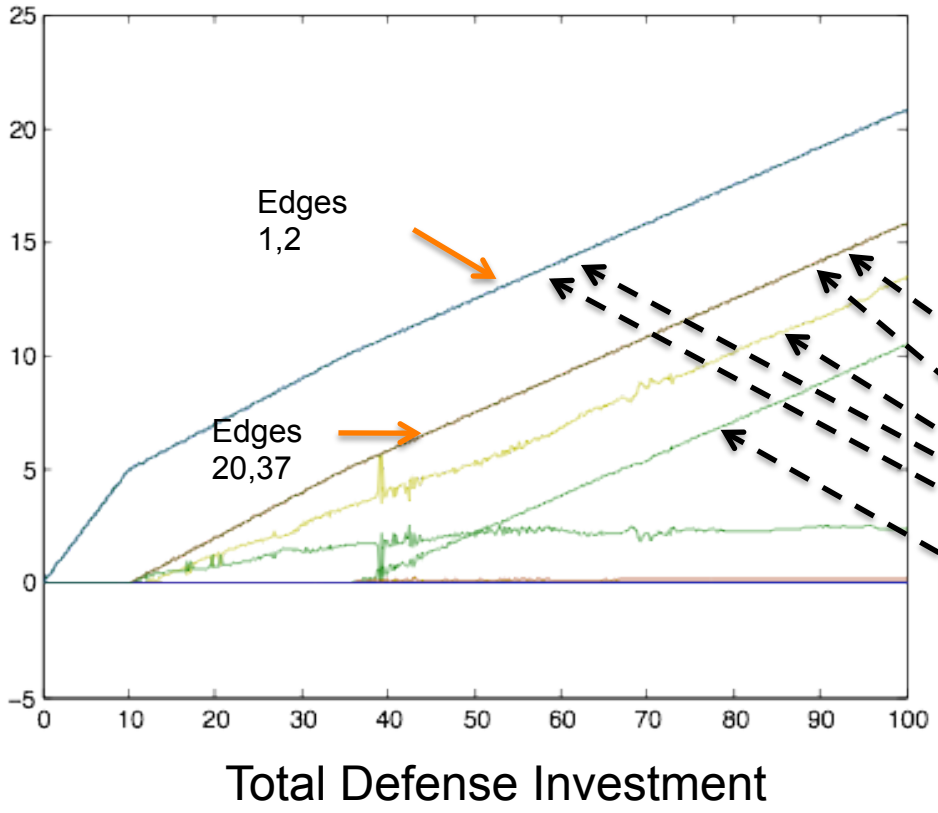
If we are given an attack graph with

- a minimal cut set that has e edges
- a large investment budget, K

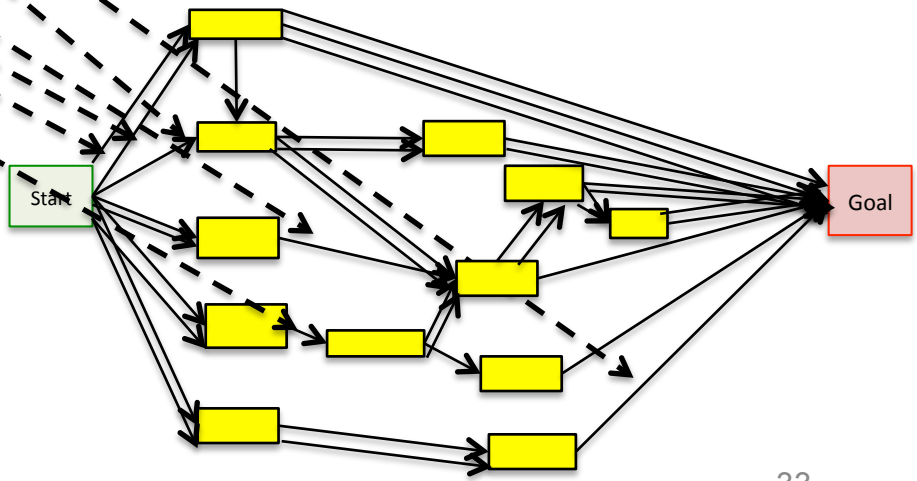
then

- the optimal budget allocation assigns ≈ K/e to each edge in the cut set and;
- the minimal cost path grows like c + K/e where c is a constant

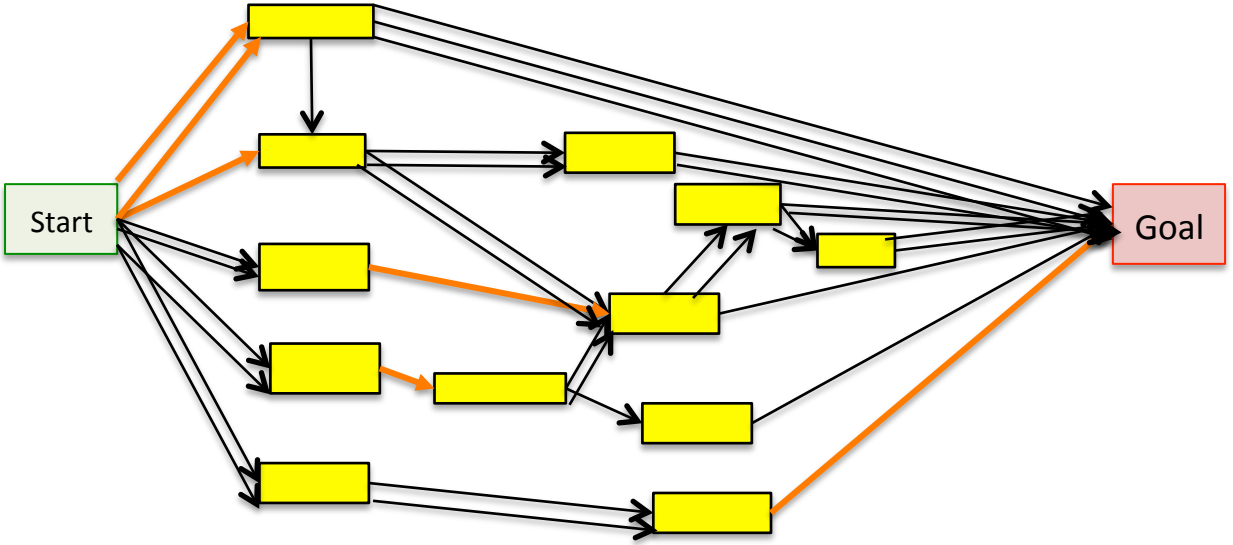
Linear Programming Results Identify High Value Protection Paths for Different Investment Levels



- Theorem states that optimal investment is eventually K/e in minimal cut set edges
- Initially, optimal investments can occur in other edges



Back to Real System



37 edges
180 paths
12 nodes
e = 6, cut set

Multiple edges mean multiple attack steps possible

Matrix M has 37 columns and 180 rows

Adversarial Dynamics Takeaways 1/2

- "Big data" needed
  - Red and blue forces' data sets are needed
  - New, *non-stationary* statistics and estimation are key
  - Adaptation, *not static equilibria*, describe "solutions"
- "Hidden data" needed
  - Need to capture **what players/agents think, not just the outcomes**
- Anticipating moves is the way to gain advantage
  - Kasparov who can think 5-6 moves ahead

References

1. Cybenko, Landwehr, "Security Analytics and Measurements", IEEE S&P , May-June 2012 http://tinyurl.com/securityanalytics
2. Bilar, Murphy, Cybenko, "Conficker Case Study", in MTD II (ed. Jajodia), 2012 http://tinyurl.com/confickerQAG
3. Saltaformaggio, Bilar "ABCD-ACP", ICCC3 NATO CCD COE, 2011 http://tinyurl.com/ICCC3
4. Stocco, Cybenko, "Inverse game theory", SPIE 8359, 2012 http://tinyurl.com/inversegame
5. Carin, Cybenko, Hughes, "Queries methodology", IEEE Computer, 2008 http://tinyurl.com/queries2008
6. Ohtsuki, Novaw, "Replicator equations", Journal of Theoretical Biology 243 (2006) 86–97 http://tinyurl.com/replicationequ

Thank you

Thank you for the kind consideration of these ideas

We are happy to answer questions / field comments ☺

Contact:

- Daniel Bilar: dbilar@acm.org
- George Cybenko: gvc@dartmouth.edu
- John Murphy: jmurphy@proquesys.com

Additional Slides

Oscillations as Manifestation of Adversarial Dynamics

- Evolution is a response to competition
- Competition exists among adversaries
- How do you know you are operating in an "adversarial" domain?
  - Oscillations of performance metrics
- Dynamics can be modeled by replicator equations
  - Typically 3rd order, non-linear (analytically difficult)
- *Inverse problem* of observing behavior and estimating parameters of replicator equation that guide behavior is tractable
- Possible to observe game play and strategy evolution and then make inferences about player's motives, costs and move options