Network design games in presence of strategic adversaries

Prithwish Basu

Lead Scientist

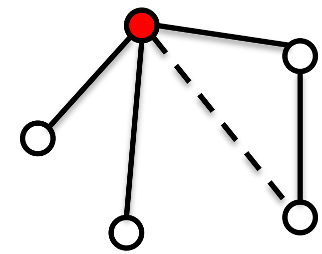
Raytheon BBN Technologies

Overview

- Network design problems
  - Designing or re-designing networks to improve desirable properties
- Adversarial models
- Focus of this talk
  - Strategic adversary
  - Non-cooperative game formulations
  - Topology sequences
  - One-shot games and Markovian variants
  - Multi-stage games
- Relevance to this workshop
  - Topology dynamics has direct impact on spread of epidemics
  - So, one could design networks for facilitating or curbing epidemics, while an adversary may want the opposite

Network design for a purpose

- Given:
  - a network (graph) **G=(V, E)**
    - **G** could have weights on edges and/or nodes
    - a property **P** defined on **G**
  - a cost budget **B**



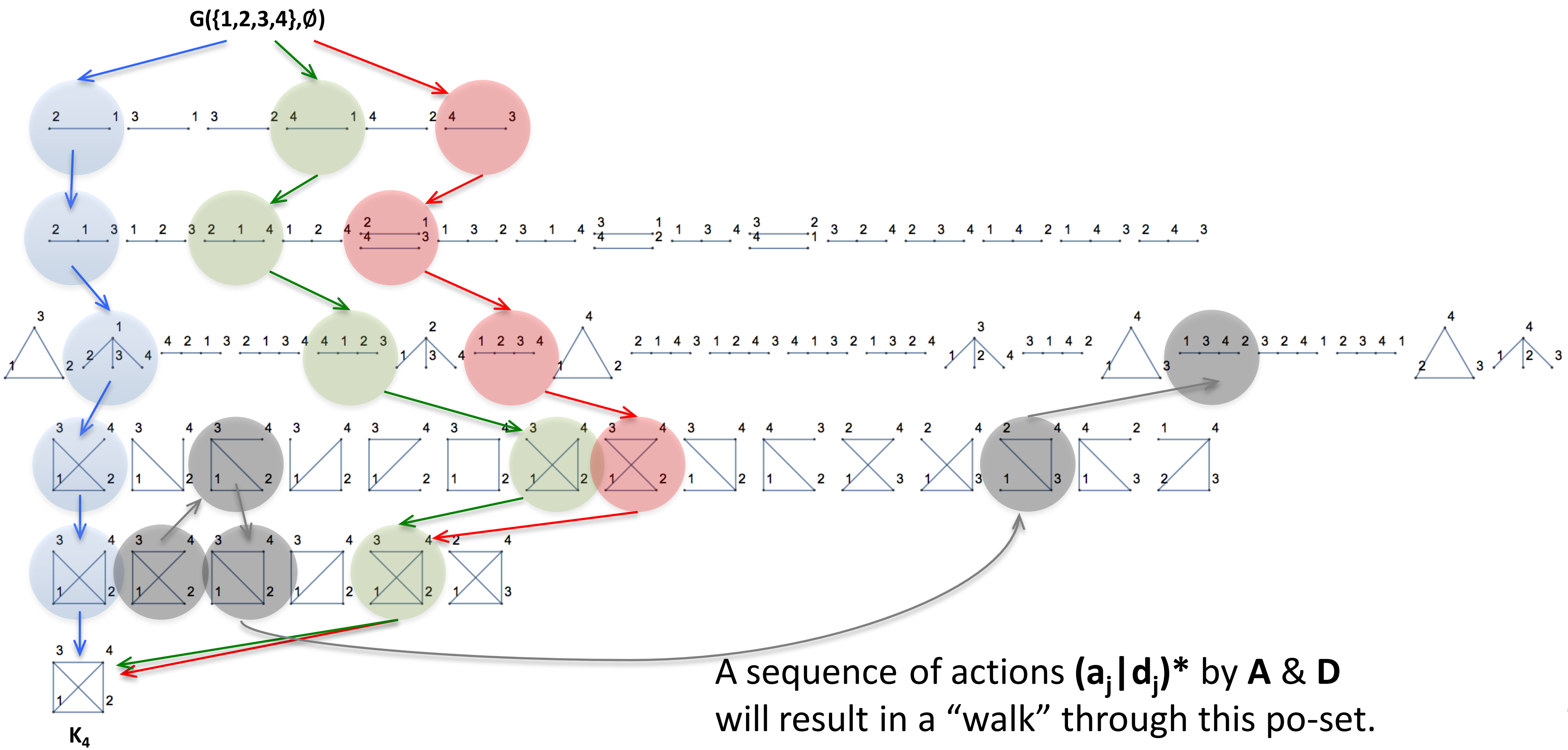
eccentricity(●) = 2 → 1

- Edge problems
  - Add **B** edges from $\mathbf{G^c = (V, K_{|V|} \setminus E)}$ to G such that **P** is minimized (or maximized)
    - **P**: *global*, e.g., diameter, average shortest path length, connectivity, etc.; or *local*, e.g., eccentricity or betweenness centrality of a node
    - Problems typically NP-complete if **B** is part of the input
- Node problems
  - If **G** has positive node weights, select **B** nodes whose weights can be reduced to **0** such that **P** is minimized
    - **P**: average latency (also NP-complete)

Adversarial action over time

- Consider an adversary (**A**)
  - Adversarial action: remove edges: $\mathbf{G_t \rightarrow_a G_{t+1} \subset G_t}$
  - Loss of edges typically results in worse value of **P**
- Network designer (**D**) has to take *action*
  - Just restore the old topology: $\mathbf{G_t \rightarrow_a G_{t+1} \rightarrow_d G_t}$
  - OR add different edges: $\mathbf{G_t \rightarrow_a G_{t+1} \rightarrow_d G'_t \neq G_t}$
- The space of all possible topologies is a *partial order (po-set)*, and **D** and **A** would bounce around that po-set

Po-set of network topologies



A sequence of actions **(a_j|d_j)*** by **A** & **D** will result in a "walk" through this po-set.

Goal: study interesting properties of this dynamical process under different adversarial models.

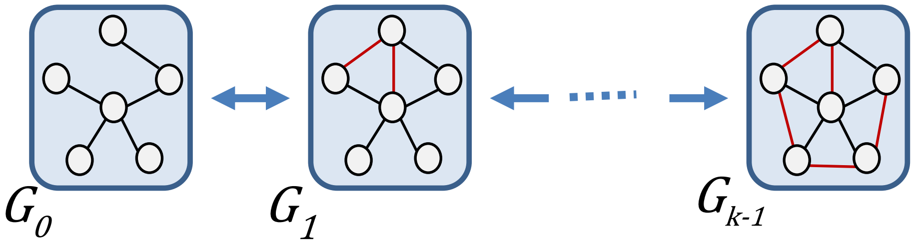
NETWORK SCIENCE CTA

Policy-compliant topologies

Raytheon BBN Technologies

Tractable case: Dynamics along a sequence of operationally allowed or **“policy-compliant”** topologies

Base topology



- Nodes for topology G_i : V_i , set of edges: E_i
- **Densification property**: $\forall i$: $V_i = V$, but $E_0 \subset E_1 \subset \ldots \subset E_{K-2} \subset E_{K-1}$
- If $\mathbf{|E_i \backslash E_{i\text{-}1}| = 1}$, it is basically a vertical path of length **K** through the po-set of topologies

Examples and Rationale

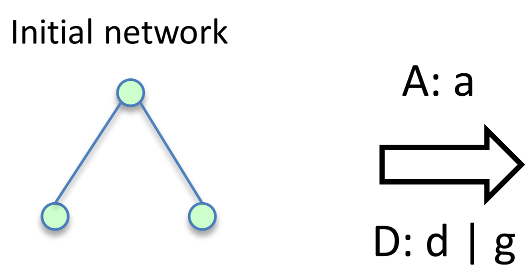
- Each edge may correspond to a new pair-wise association, e.g., shared key
- The order of associations is important since dependencies may be involved
- If two managers M_1 and M_2 are given a shared key, and their employees S_1 and S_2 are too, removal of the M_1—M_2 relationship would invalidate S_1—S_2 relationship as well
- Thus, attack on edge ***j*** in state $\mathbf{G_{K\text{-}1}}$ would result in its removal and “backtracking” to the best policy compliant topology

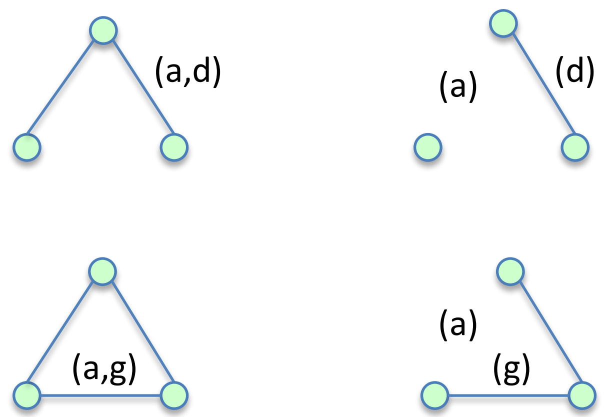
Benign adversary

- Benign adversary
  - Attacks according following some model (e.g., at random locations) and incurs zero cost
  - Examples: wireless interference, thermal noise
  - Actions not in step with that of network designer **(D)**
  - **D** wants to optimize a given property **P** and incurs action costs (to **add** / **edit** / **maintain** edges)
- Solution approach
  - Stochastic Dynamic Programming but concentrate on instantaneous states to avoid dimensionality curse
  - This yields a *modified myopic policy*
  - *E. N. Ciftcioglu, K. S. Chan, A. Swami, D. H. Cansever and P. Basu, "Topology Control for Time-Varying Contested Environments", MILCOM 2015.*

Focus of talk: strategic adversary

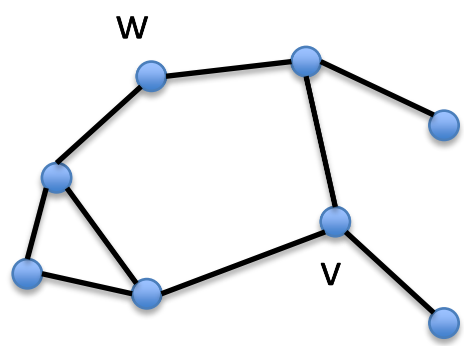
- Strategic adversary **(A)**
  - Observes the network and attacks where it hurts the most
  - Examples: cyber attacks
  - **D** and **A** incur costs for actions defend (**d**), grow (**g**), or attack (**a**)
  - Actions occur *simultaneously* with that of network designer **(D)**
  - Solution approach: model the scenario as a *2-player one-shot non-cooperative game*
- Rules of the game (when not restricted by a policy compliant sequence)





Monitor placement game (nodes)

First, consider a related framework where actions are on nodes



D can place monitor at any node
A can attack monitor port at any node

- Where to place a monitor/controller in presence of a strategic adversary (**A**)?
- Optimization metric: **eccentricity** of monitor node **v**
  - $\mathbf{e_v}$: max {shortest paths from **v**}

– If **D** places monitor on node **v** *and* **A** guesses correctly and attacks **v**, then
  - Utility, **U = 0**

– If **D** places monitor on **v** *and* **A** guesses wrongly and attacks the monitor port of node **w ≠ v**, then
  - Utility, $\mathbf{U = 1/e_v}$

Non-cooperative game (**A** vs. **D**)

- Consider probabilistic strategies for
  - Placement (by **D**): $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$
  - Attack (by **A**): $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_n)$
- Since $\mathbf{e}_v \geq \mathbf{1}$, $\mathbf{0} \leq \mathbf{U} \leq \mathbf{1}$
  - Low U: bad; High U: good
- Expected utility: quadratic form

$$E[U] = p^T M \, q$$

$$M = \begin{bmatrix} 0 & \frac{1}{e_1} & \frac{1}{e_1} & \frac{1}{e_1} \\ \frac{1}{e_2} & 0 & \frac{1}{e_2} & \frac{1}{e_2} \\ \frac{1}{e_3} & \frac{1}{e_3} & 0 & \frac{1}{e_3} \\ \frac{1}{e_n} & \frac{1}{e_n} & \frac{1}{e_n} & 0 \end{bmatrix}$$

Solution to the Matrix Game

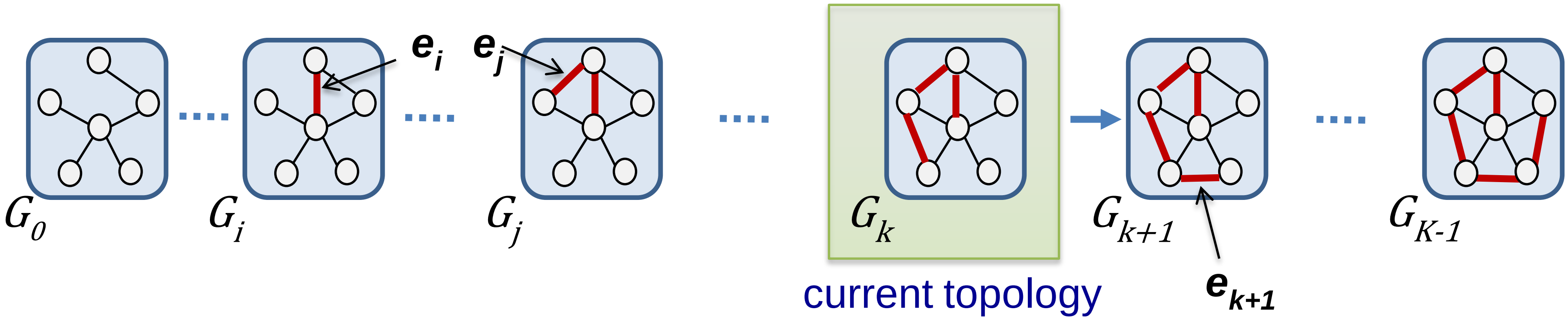
- One-shot 2-player zero-sum bimatrix game with standard assumptions of rationality, knowledge etc.
  - Mixed Nash equilibrium must exist
- Expected utility: $E[U] = V = \sum_{j=1}^{n} \sum_{i=1}^{n} p_i M_{ij} q_j$
- ***M*** has special structure => solvable in closed form by using the **principle of indifference** $\sum_{i=1}^{n} p_i M_{ij} \geq V$
- Equilibrium solution structure
  - Placement probabilities, $p_i^* = \frac{e_i}{\sum e_i}$ (Tends to place at high eccentricity nodes!)
  - Attack probabilities, $q_j^* = 1 - \frac{(n-1)\, e_j}{\sum e_j}$ (Tends to attack low eccentricity nodes)
  - $E[U^*] = V = \frac{n-1}{\sum e_i}$ (Utility at Nash Equilibrium)

NETWORK SCIENCE CTA

Policy-compliant topology game

Raytheon BBN Technologies

[Ciftcioglu, Pal, Chan, Cansever, Swami, Singh, and Basu, WiOpt 2016]



At topology state **k**, **D** and **A** act simultaneously:

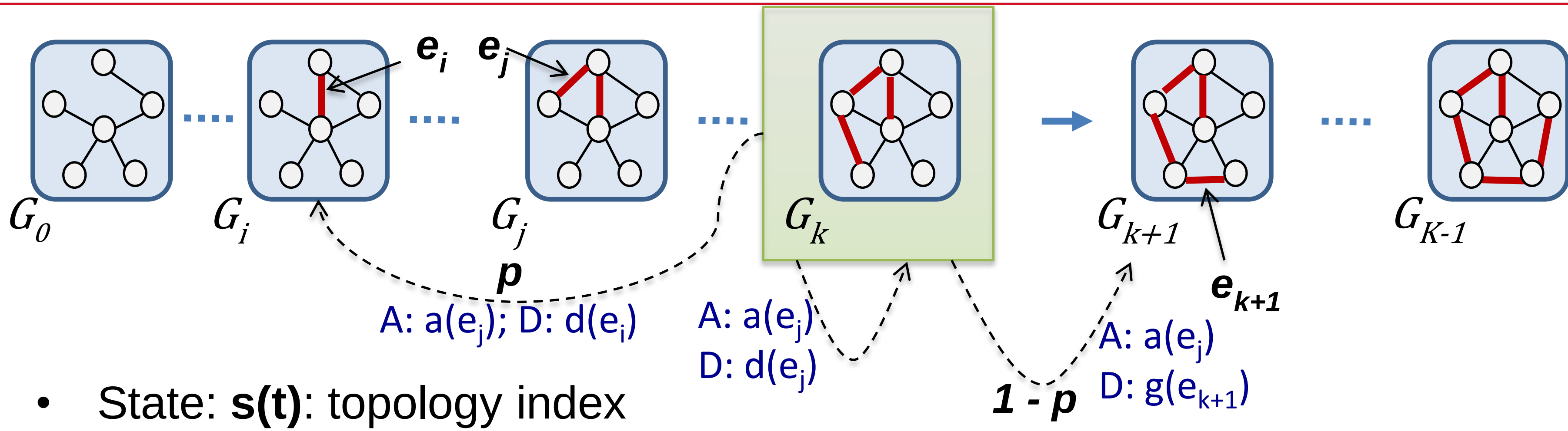
Designer Action: **D** either chooses to protect one of the edges, or further grow the network by adding a edge, either:

- Defend an existing edge $\boldsymbol{e_i}$, or
- Try to grow the network by adding edge $\boldsymbol{e_{k+1}}$

Adversarial Action: **A** intelligently tries to disrupt network functionality by attacking edges, either:

- Attack an existing edge $\boldsymbol{e_j}$
- Attack an “anticipated” edge $\boldsymbol{e_{k+1}}$

State Transitions



- State: **s(t)**: topology index
- Attack success probability ***p*** (results in state transitions)
- If an edge is not defended, **A** disrupts it with probability ***p***
  - If attack successful, D has to *backtrack* to the allowed topology that can be formed by the remaining edges
  - If attack unsuccessful, network can grow depending on D's strategy.

$$s(t+1) = \begin{cases} a(t) - 1, & \text{w.p. } p, \text{if } a(t) \neq d(t) \\ s(t), & \text{if } a(t) = d(t), \text{or w.p. } (1-p) \text{ if } a(t) \neq d(t) \\ s(t) + 1, & \text{w.p. } 1 - p, \text{if } d(t) = s(t) + 1 \end{cases}$$

Payoffs and Costs

Designer: (d_k)

- Cost of defending existing edge: **δ**
- Cost for adding a new edge: **γ**

Typical Assumption: (**$\delta < \gamma$**): growing edges more costly

Adversary: (z_k)

- Cost of attacking existing edge: **β**
- Cost for attacking an anticipated edge: **α**

Typical Assumption: (**$\beta < \alpha$**): existing edges more established

Overall utility: Network property cost (g_k) + Own operational costs:

Designer: *minimize* $g_k + d_k \equiv \boxed{\textit{maximize} - g_k - d_k}$

Adversary: $\boxed{\text{maximize } g_k - z_k}$

For many results, we assume **$\delta = \gamma = \beta = \alpha = 0$** => **zero-sum** game

Properties of Nash Equilibria

A: Attacked edge ID   Game matrix at state ***k***)

D

Defend

Grow

$d \backslash a$	1	$\dots$	k	$k+1$
1	$(-g_k - \delta, g_k - \beta)$	$\dots$	$(-pg_{k-1} - (1-p)g_k - \delta, pg_{k-1} + (1-p)g_k - \beta)$	$(-g_k - \delta, g_k - \alpha)$
2	$(-pg_0 - (1-p)g_k - \delta, pg_0 + (1-p)g_k - \beta)$	$\dots$	$(-pg_{k-1} - (1-p)g_k - \delta, pg_{k-1} + (1-p)g_k - \beta)$	$(-g_k - \delta, g_k - \alpha)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
k	$(-pg_0 - (1-p)g_k - \delta, pg_0 + (1-p)g_k - \beta)$	$\dots$	$(-g_k - \delta, g_k - \beta)$	$(-g_k - \delta, g_k - \alpha)$
$k+1$	$(-pg_0 - (1-p)g_{k+1} - \gamma, pg_0 + (1-p)g_{k+1} - \beta)$	$\dots$	$(-pg_{k-1} - (1-p)g_{k+1} - \gamma, pg_{k-1} + (1-p)g_{k+1} - \beta)$	$(-pg_k - (1-p)g_{k+1} - \gamma, pg_k + (1-p)g_{k+1} - \alpha)$

(***g***g_k: Network property cost at topology state ***k***)

Game does not possess pure-strategy Nash equilibrium by inspection unless special conditions where ***p*** very low:

- Strategy of growth optimal if $p < \frac{g_k - g_{k+1}}{g_0 - g_{k+1}}$
- If g_k concave decreasing, growth optimal if $p < \frac{1}{k+1}$
- If g_k convex decreasing, no pure strategy by inspection if $p > \frac{1}{k+1}$

In general, both D & A play *mixed (probabilistic) strategies*

Decisions to network evolution

- Designer and attacker play mixed (probabilistic) strategies for choosing edges
- Result: stochastic topology dynamics
  - Due to randomness in actions, and attack success
- Can be modeled by a **Markov game**
  - What are the structural properties of mixed strategies?
  - What are the state transition probabilities? (Computable from game rules and strategy profiles)
  - What is the steady state probability of being in each topology?

Incentives for Designer & Adversary

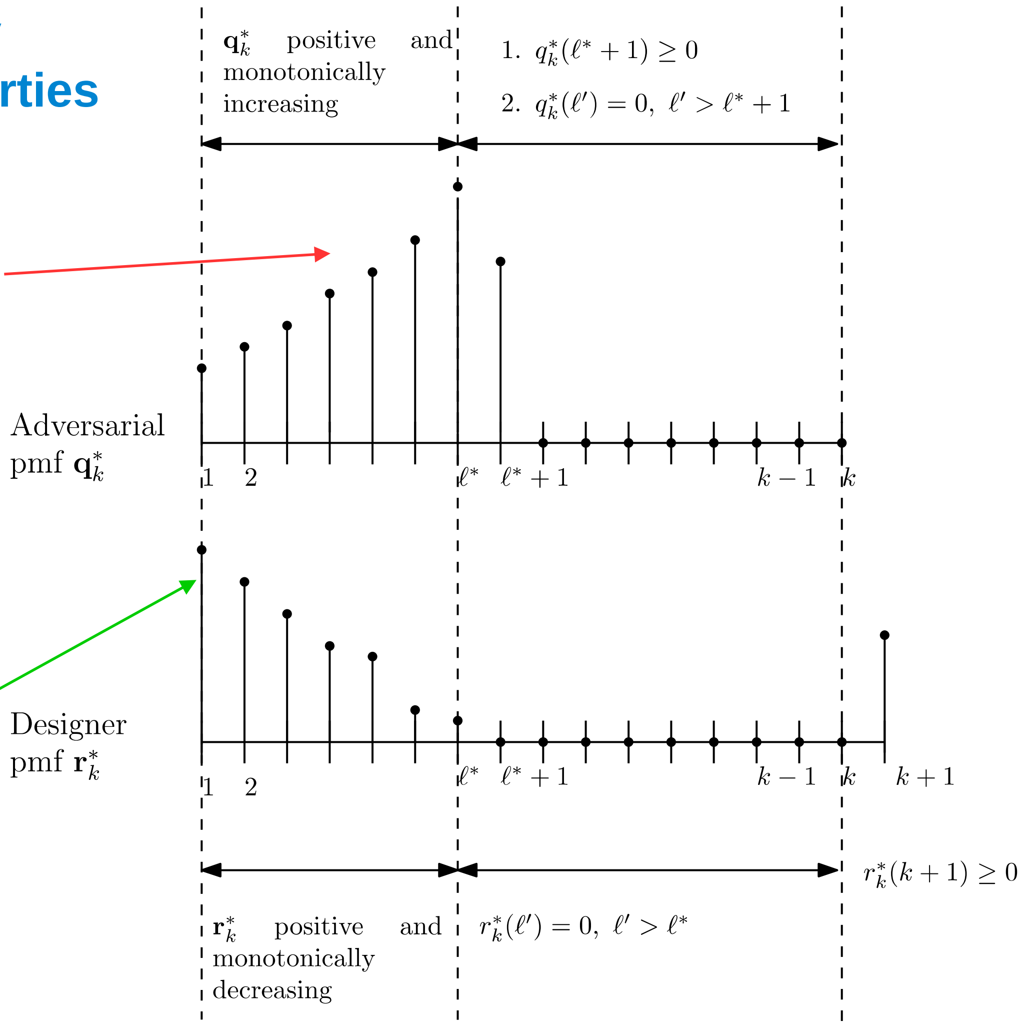
- Initial intuition
  - Adversary: targets important edges to inflict maximum damage, and
  - Designer: prioritizes defense of important edges
- However, two phenomena
  - Adversary's view: Since **D** might defend the most crucial edges, any attack on those edges might be neutralized, therefore **A** shifts focus on attacking “important” edges but not the “most important” ones
  - Designers view: If ***p*** is small, why not take chances and try to grow the network?

Properties of Mixed Nash Strategies

Equilibria exhibit nice distributional monotonicity for monotonic graph properties

Adversary attacks less important links with greater probability to avoid hitting a defense wall!

Designer acts as expected, prioritizes more important links to avoid deep backtracking



Strategy Probabilities to Transition Probabilities

Obtain state transition probabilities $\gamma_{k,j}$ from state k to state j as a function of mixed strategy probabilities:
Designer: $(r^*_k(1), \ldots, r^*_k(k), r^*_k(k+1))$
Adversary: $(q^*_k(1), \ldots, q^*_k(k), q^*_k(k+1))$
and attack success probability p:

$$\gamma_{k,0} = q^*_k(1)(1 - r^*_k(1))p$$

Degrading to base topology

$$\gamma_{k,k+1} = (r^*_k(k+1))(1-p)$$

Growing to next topology

$$\gamma_{k,j} = q^*_k(j+1)(1 - q^*_k(j+1))p$$

Backtracking to topology j from k, j<k

$$\gamma_{k,k} = \sum_{j=1}^{k} [r^*_k(j)q^*_k(j) + (1-p)(1 - r^*_k(j)q^*_k(j))] + r^*_k(k+1)q^*_k(k+1)p.$$

Staying at the same topology

Steady State Probabilities

Once mixed strategies and resulting state transition probabilities found, construct State transition matrix

$$P = \begin{pmatrix} \gamma_{0,0} & \gamma_{0,1} & 0 & \cdots & 0 & 0 \\ \gamma_{1,0} & \gamma_{1,1} & \gamma_{1,2} & \cdots & 0 & 0 \\ \gamma_{2,0} & \gamma_{2,1} & \gamma_{2,2} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \gamma_{k-1,0} & \gamma_{k-1,1} & \gamma_{k-1,2} & \cdots & \gamma_{k-1,k-1} & \gamma_{k-1,k} \\ \gamma_{k,0} & \gamma_{k,1} & \gamma_{k,2} & \cdots & \gamma_{k,k-1} & \gamma_{k,k} \end{pmatrix}$$

Balance equations and equilibrium distribution found using

$$\pi P = \pi$$

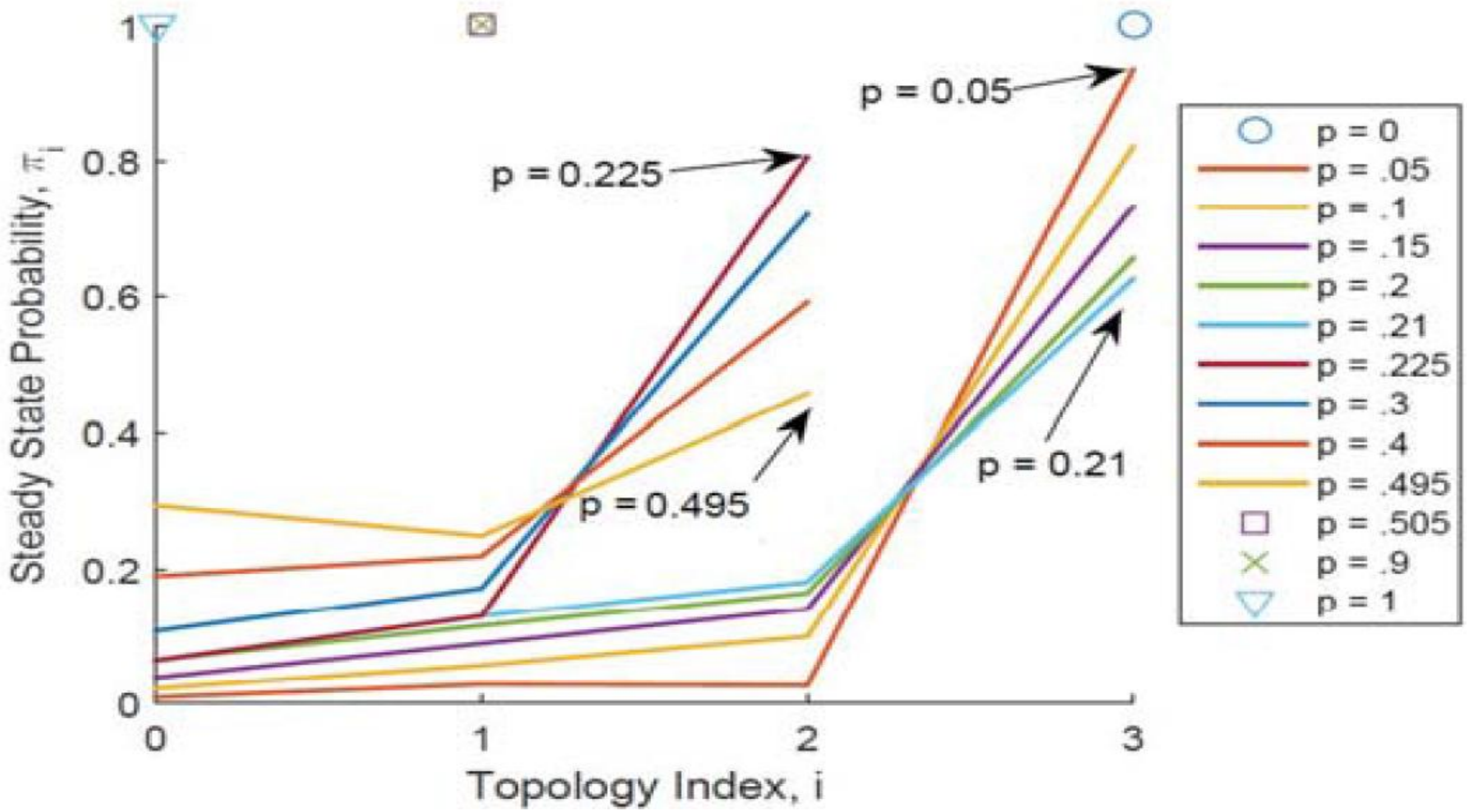
Along with

$$\sum_{i=0}^{K} \pi_j = 1.$$

Numerical Results
Varying Attack Success Probability

Network property: Harmonic mean of path lengths

No operational costs, start from base topology

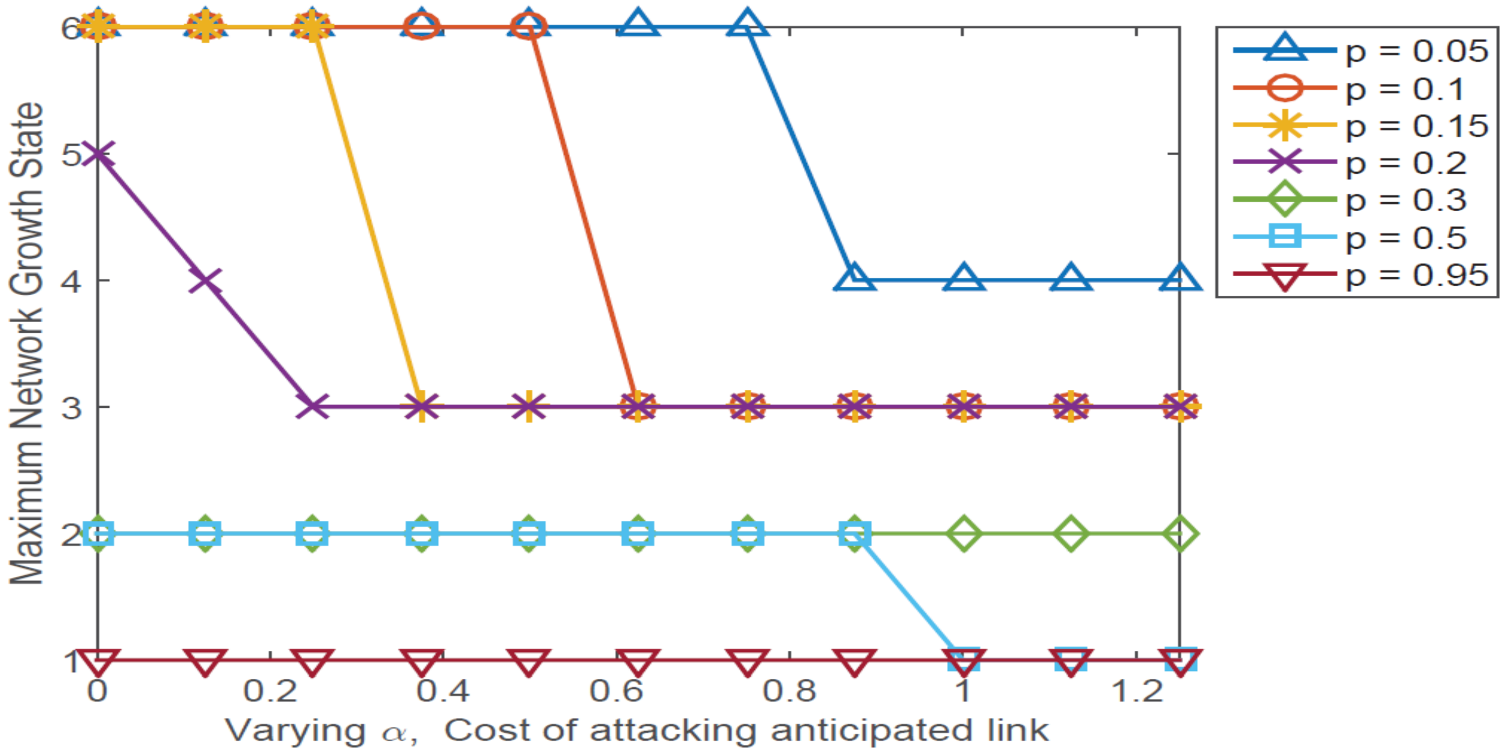


Starting from G_0, the network can grow to higher states for lower ***p***

Numerical Results

Effect of Operational Costs

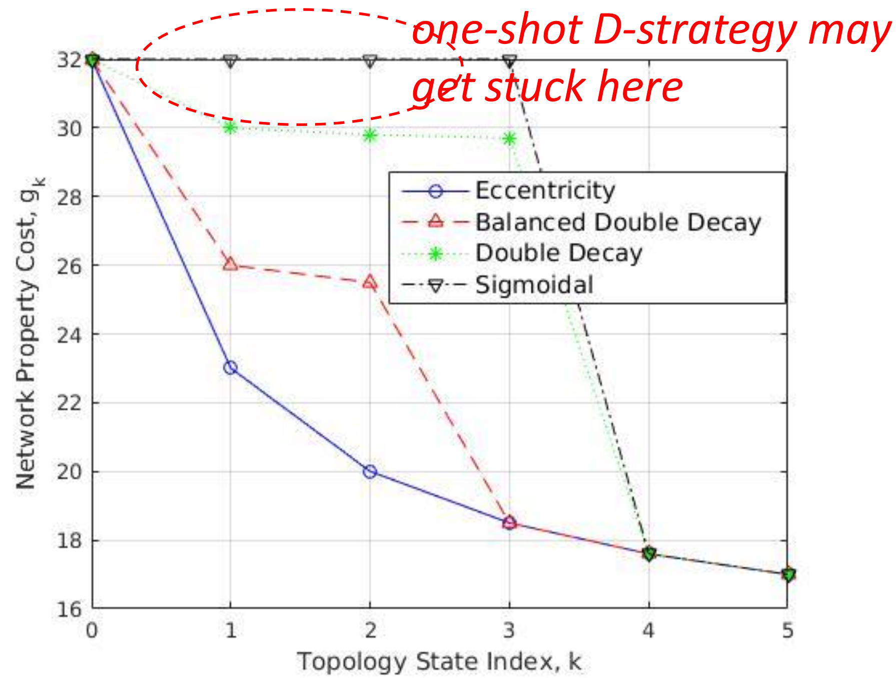
Network property: Eccentricity



When the adversary is capable of performing with lower operational costs α, the network can eventually evolve to larger sizes!

Beyond one-shot games

- So far, **D** and **A** have played repeated instances of one-shot games
- Being more adventurous is ideal sometimes
  - e.g., the $\boldsymbol{g_k}$ functions can have complex structures that result in suboptimal behavior



- Play a *multi-stage* game
  - Maximize a discounted sum of rewards over a time horizon
  - With no adversary – this is the MDP framework
  - With adversary – multi-stage Markov game

Multi-stage games

- Value functions of D & A consider potential future rewards:

$$V_D(k,\mathbf{r},\mathbf{q}) = \sum_{t=0}^{\infty} \gamma^t E[y_t^D|\mathbf{r},\mathbf{q},k]$$

- Mixed Nash for this game exhibits similar monotonicity properties as the one-shot game
- Algorithms from Markov-games literature
  - **Q-learning** $$\mathcal{Q}_d^*(k,d,a) = U_1^k(d,a) + \gamma \sum_{k'\in S} T(k'|k,a,d)V_d(k',\mathbf{r}^*,\mathbf{q}^*)$$

    Iterative: $$\mathcal{Q}_d^{i+1}(k,d,a) = (1-\alpha)\mathcal{Q}_d^i(k,d,a) + \alpha(-g_{k'} + \gamma V_d^i(k'))$$

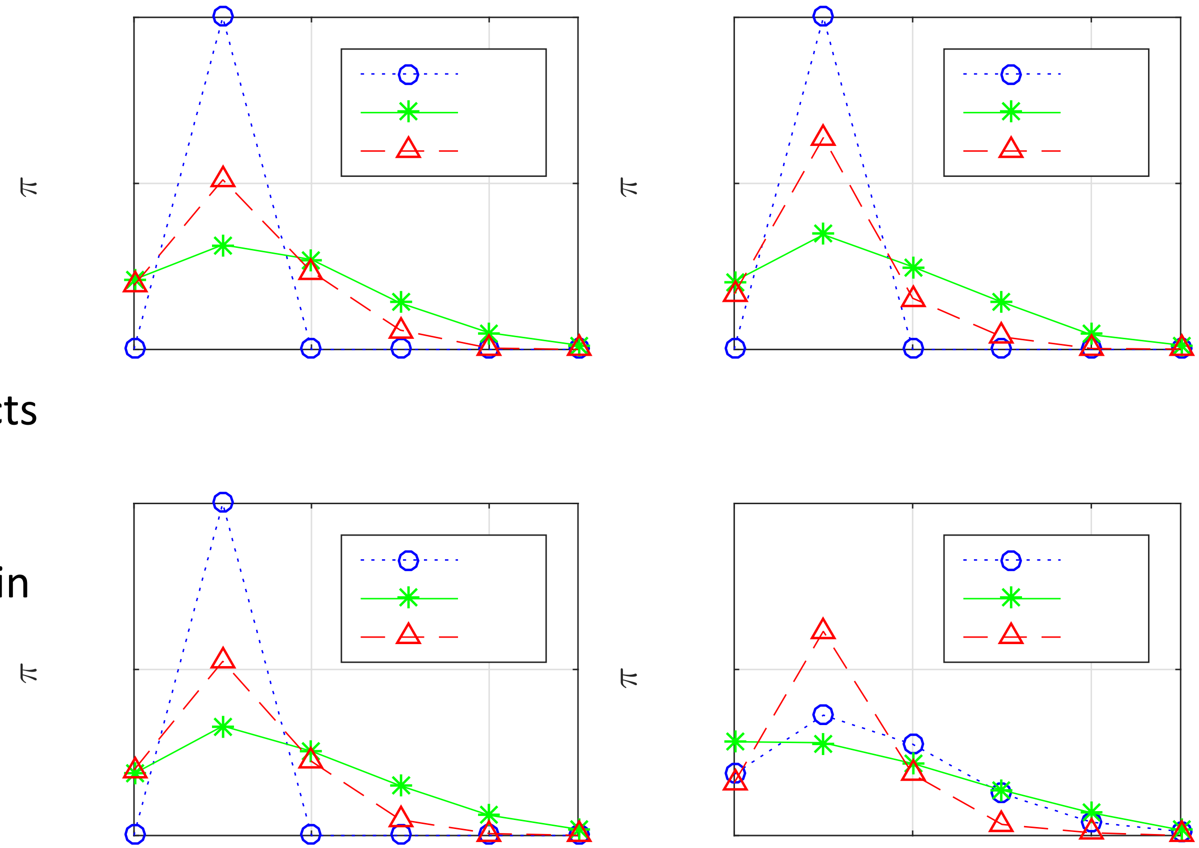
    $$V_d^i(k') = \mathbf{r}_i^*(k')\mathcal{Q}_d^i(k')\mathbf{q}_i^*(k')$$

    learning rate
  - **Rollout policies**
    - Consider all one-step (a, d | g) action pairs and simulate further actions (Monte Carlo) using base policies: then update the game matrix entries
    - This is less computationally intensive than Q-learning

Numerical results

Steady state topologies

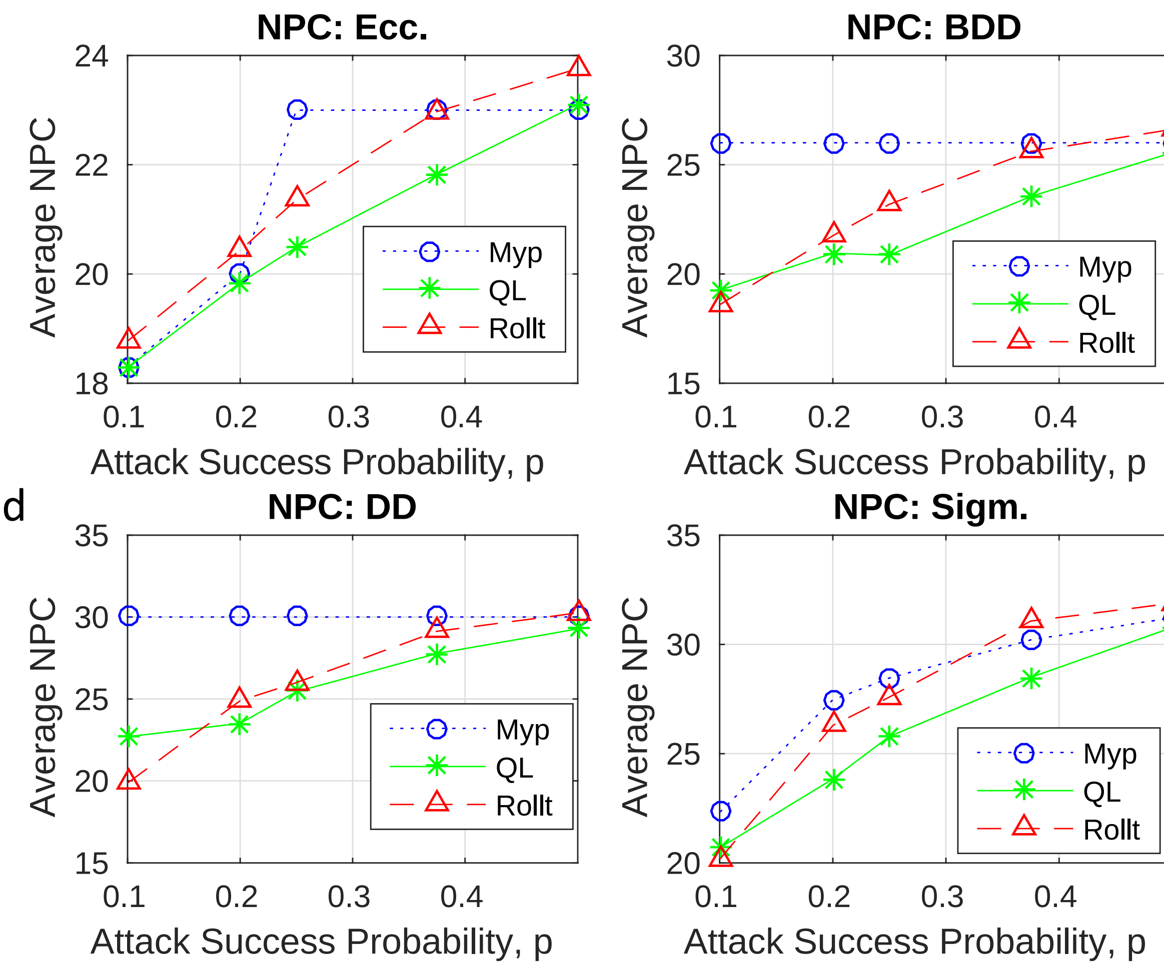
p = 0.5



The *exploration step* of **Q-Learning** randomly selects growth strategies even at high **k**, when the risk of backtracking outweighs gain from growth.

Q-Learning is able to take the network to higher states than **Rollout** and **one-shot**

Numerical results
Time-averaged network cost g_k



Sometimes at high ***p***, the **one-shot** policy does well compared to **Q-Learning** and **Rollout**, because it tends to protect from backtracking all the way to G_0.

Q-Learning is generally the best policy in the mix

Ongoing research directions

- Relax assumptions about
  - complete knowledge of the network state
  - knowledge of the payoff structures
  - knowledge of others' actions and resources
- Gain fundamental understanding of co-evolution of networks in adversarial settings resulting from
  - interaction between multiple networks
  - interaction between network structure and information flow
- Decentralized behavior in adversarial settings
  - multi-party games, coalition formation etc.

Acknowledgements

Collaborators

- Ananthram Swami & Kevin Chan (US Army Research Labs)
- Ertugrul Ciftcioglu (IBM Research & ARL)
- Derya Cansever (US CERDEC)
- Siddharth Pal (BBN)
- Ambuj Singh (UCSB)
- Christos Faloutsos (CMU)

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053 (the ARL Network Science CTA). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.