DYNAMIC STOCHASTIC NETWORK INTERDICTION GAMES

David Castañón

work with Dr. J. Zheng, now at Google
A bit of background

- Demos and I were graduate students together at MIT (1970s) in ESL (becomes LIDS in 1978)
- Control theory evolving from servomechanisms to broader classes of systems in late 1960s-1970s
  - Economics, Transportation, Power Systems, military systems, organizations, manufacturing, communications, ...
    - Large scale systems, multiple agents, distributed information, multiple objectives, ...
  - Extensions of fundamental theories were explored
    - Maximum principle, stochastic control, linear-quadratic-Gaussian optimization, dynamic programming, ...
- Many limitations were discovered!
  - Witsenhausen counterexample, non-nested information, ...
COMMUNICATION IN DECENTRALIZED CONTROL

by

Demosthenis Teneketzis
B.S. University of Patras
Patras, Greece
1974
S.M. Massachusetts Institute of Technology
1976
E.E. Massachusetts Institute of Technology
1977

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1979
(e. February 1980)
More background

- Demos was very interested in the overlap of communications and distributed control
  - PhD Thesis on Real-time rate distortion problems
    - Encoder and Decoder as decentralized controllers with different information
    - Focus on bounds for nonclassical information problems
  - Demos subsequently joins industry (ALPHATECH)
    - Continues research on large scale systems with distributed information
    - Interesting fundamental work on team problems, games, consensus (real Bayesian, not the simpler linear one), ...
      - Decentralized Wald & Quickest detection, asymptotic agreement, subjective games, distributed estimation, ...
Story ends well

- Demos leaves ALPHATECH, joins U. Michigan, finds a welcoming Department
  - Precursor to this wonderful event
  - Example for others to follow: DC, Ahmet Tewfik, ...

- Continues research directions
  - Adds strong academic components of teaching and mentoring to his excellent research

- Avoids the curse of administration!
  - Manages to stay on the academic track
Problem in this Talk: Network Interdiction

- Game between an intelligent network and an attacker
  - Network knows how to adapt operations to network conditions
  - Attacker can degrade parts of the network, anticipating the network’s reaction
  - Zero-sum: Attacker wants to maximally degrade network performance

- Many applications:
  - Power distribution, transportation, communications, drug smuggling, cybersecurity, border patrols, ...
Examples:

- Removal of links, switches in transmission networks with adaptive routing
- Degradation of roads to limit throughput
- Deployment of air marshals on flights
- Deployment of sensors for border monitoring
- Saturation of switches in comms networks through cyber attacks
- ...

Previous Work

- One-stage combinatorial optimization techniques
  - Early work (McMasters-Mustin ’70, Ghare et al ’71, Fulkerson-Harding ’77, Golden ’78): interdiction of flow networks
  - Wood ‘93, Whiteman ‘99: min-max Stackelberg network game
  - Cormican et al ‘98, others: uncertain outcomes in attack effectiveness
    → Stochastic min-max Stackelberg game with perfect information
  - Zenklusen ’08, Akgun et al ‘11: multiterminal network interdiction
  - ZC ‘12: Faster algorithms for single stage stochastic network interdiction, extension to dynamic attack with success feedback

- Mostly focused on one stage per player
  - Act – React
  - Stochastic effects present, but fully observed with perfect information
  - Computation algorithms
Application: Power Systems

- Green diamonds are sinks with demands
- Blue squares are generators with capacity
- Arcs have maximum power flow capacity
- Goal: Maximally disrupt service by disrupting arcs
Stochastic network interdiction

- Directed network, nodes $N$, arcs $A$, with source node $s$, terminal node $t$ that wants to send max flow
- Player Y has a budget to attack arcs
  - Attacked arc $(i,j)$ removed with probability $p_{ij}$
- Outcome of attack on $(i,j)$ is $w_{ij} = 0$ (arc survives) or 1
- Network response: max flow given surviving arcs

\[
h(\omega) \triangleq \max_x x_{ts}
\]

s.t. \[
\sum_{(n,i) \in \overline{A}} x_{ni} - \sum_{(j,n) \in \overline{A}} x_{jn} = 0, \quad \forall n \in N
\]
\[
x_{ij} \leq u_{ij}(1 - \omega_{ij}), \quad \forall (i,j) \in A
\]
\[
x_{ij} \geq 0, \quad \forall (i,j) \in \overline{A}.
\]
Attacker’s Problem

- Budget $R$, cost of attacking arc $(i,j)$ as $c_{ij}$
- Notation: $l_{ij}$ is arc outcomes if attacked:

$$J \triangleq \min_{\gamma} \sum_{I \in \Omega} P(I) \ h(I \cdot \gamma)$$

$$s.t. \sum_{ij} c_{ij} \gamma_{ij} \leq R; \ \gamma_{ij} \in \{0, 1\}$$

$$P(I) = \prod_{(i,j) \in A} (1 - p_{ij})^{1-I_{ij}} p_{ij}^{I_{ij}}$$

- Issues: Exponential # terms, integer program, $h$ concave in relaxed variables, ...
Alternative formulation: Penalty max flow

\[
g(\omega) \triangleq \max_x x_{ts} - \sum_{(i,j) \in A} \omega_{ij} x_{ij}
\]

\[
\sum_{(n,i) \in \overline{A}} x_{ni} - \sum_{(j,n) \in \overline{A}} x_{jn} = 0, \forall n \in N
\]

\[
x_{ij} \leq u_{ij}, \forall (i, j) \in A
\]

\[
x_{ij} \geq 0, \forall (i, j) \in \overline{A}
\]

- \( h(w) = g(w) \), and \( g(w) \) is convex on \([0,1]^{|A|}\)
- Relaxations now feasible, yield bounds
Useful Bounds

- Jensen's inequality:
  \[
  \min_{\gamma \in \{0,1\}} \sum_{i,j} c_{ij} \gamma_{ij} \leq R \quad \min_{\sum_{i,j} c_{ij} \gamma_{ij} \leq R} E_I[g(I \cdot \gamma)] \geq \min_{\gamma \in [0,1]} \sum_{i,j} c_{ij} \gamma_{ij} \leq R \quad E_I[g(I \cdot \gamma)] \\
  \geq \min_{\gamma \in [0,1]} \sum_{i,j} c_{ij} \gamma_{ij} \leq R \quad g(E[I] \cdot \gamma)
  \]

- Bound: single max-flow problem

- Tighter bounds: Let $F$ be a partition of $\{0,1\}^{|A|}$.
  \[
  \min_{\gamma \in [0,1]} \sum_{i,j} c_{ij} \gamma_{ij} \leq R \quad \min_{\sum_{i,j} c_{ij} \gamma_{ij} \leq R} E_I[g(I \cdot \gamma)] \geq \min_{\gamma \in [0,1]} \sum_{f \in F} P(f) g(E[I \cdot f] \cdot \gamma) \\
  \]

  \[\Rightarrow\] can use a filtration of bounds, combine with B&B
Results

- Comparisons on IEEE BUS 300 example
  - 409 arcs, 300 nodes, 33 sources, 36 sinks
  - Compare our algorithm (MBB) with two sampling-based approximate methods (SAM, SAA) from recent literature proposed as fastest
    - Exact solution (MBB) vs approximate (SAM, SAA)

No. of integer solutions = \(6 \times 10^{17}\)
Multi-Stage Network Interdiction

- Consider case where attacker can degrade network over multiple stages
  - Network observes outcomes of attack, but attacker may not
  - Attacker may not know full extent of network – needs to observe network usage of arcs

- Result: Dynamic zero-sum stochastic game with imperfect information of network state
  - Network has perfect information, attacker imperfect
  - Nested information pattern
Example

- Simple game: two arcs, two stages
  - Stage 1: Attacker chooses 1 arc to attack
  - If attack fails, network can choose whether to send flow and reveal the arc existence, or take a loss by not using the arc
  - Stage 2: Attacker chooses which potential arc to attack based on imperfect knowledge

- Extensive form including information sets per stage

Outcomes with payoffs
Formulation: Zero sum game

- Players X (max), Y (min)
- Network state at stage $t$: $s^t$
- Order: At stage $t$, X moves (attack), then Y moves, then nature
  - $a^t$: X; $n^t$: Nature; $d^t$: Y
  - Play of the game: sequence $\sigma^T = [s^0, a^1, d^1, s^1, ..., a^T, d^T, s^T]$
  - X information set at stage $t$: $I^{X,t,\sigma} = [a^1, d^1, a^2, d^2, ..., a^{t-1}, d^{t-1}]$
  - Y information set: $I^{Y,t,\sigma} = [s^0, a^1, d^1, s^1, ..., a^{t-1}, d^{t-1}, s^{t-1}, a^t]$
- Nested information: for each $t$, $I^{X,t,\sigma}$ is subset of $I^{Y,t,\sigma}$
- $\beta$ is prior distribution on $s^0$ for player X
Key Issue: How to Solve

- **Standard approach**
  - Finite # pure strategies: information sets into admissible decisions
  - Mixed strategies: random combinations of pure strategies
  - In normal form, bilinear min-max problem for solution in terms of mixed strategies
    \[
    \min_{\mu} \max_{\nu} \mu^T C \nu
    \]
  - Linear program using duality: min-max to min-min
- **Difficulty:** # pure strategies: \(O(e^{\# \text{ info sets}})\),
  # info sets = \(O(e^{\# \text{ stages}})\)
Better approach: Behavior Strategies

- Perfect recall $\Rightarrow$ equivalent solution – but performance no longer bilinear in $X$, $Y$ behavior strategies
- Alternative: behavior to sequence strategies (von Stengel ’96)

\[
\begin{align*}
\text{MINMAX} & \quad \begin{cases}
\min_{\gamma, \theta} & e' \theta \\
s.t. & A\gamma \leq E' \theta \\
& F\gamma = f \\
& \gamma \geq 0
\end{cases} \\
\text{MAXMIN} & \quad \begin{cases}
\max_{\beta, \phi} & f' \phi \\
s.t. & A' \beta \geq F' \phi \\
& E\beta = e \\
& \beta \geq 0
\end{cases}
\end{align*}
\]

- LP duality guarantees saddle point equilibria
- # variables now linear in # nodes in extensive form $\Rightarrow$ exponential in # stages
  - A “less” large LP: useful for small problems
- Can we do better exploiting information structure?
Structure: Subgame Decomposition

- Decompose at information sets of attacker
  - Rooted at an information sets of $X$, not singletons
  - All descendants of information set are in subgame graph
  - Assume probability distribution $\beta$ on the root information set
  - Subgame $G(I^{X,t}, \beta^{(X,t)})$ corresponding to subgame on information set $I^{X,t}$ with initial beliefs on states $\beta^{(X,t)}$. 

![Diagram showing subgame decomposition with nodes and edges representing information sets and actions.](image-url)
Analysis

- **Theorem:** Let $V(\beta)$ be the optimal value of any subgame with probability distribution $\beta$ on the root state. Then, there is a finite set of vectors $Q$, s.t.

$$V(\beta) = \max_{q \in Q} \beta \cdot q$$

$\Leftarrow$ LP duality (finite number of extrema in dual polytope)

- **Lemma:** Consider the saddle point behavior strategies $u^*$, $v^*$ for players $X, Y$ in the full game. Then, the restriction of $(u^*, v^*)$ to the subgame at $l(X, 2)$ are saddle point strategies in the subgame.
Induction: Base level

- Lowest level: Root nodes with values
  - Probability $\beta$ on next highest information set of $X$
  - Payoff for $X$ on given action $a$ from $I^{X,T}$ is linear in $\beta$

- Solve for saddle point value at top level as a function of $\beta$
  - Using linear programming approach
    - Representation result: Value can be represented as $\max_{v \in Q(I^{X,T})} <\beta, v>$
    - Follows from linear program duality: only a finite number of extrema in dual feasible polytope
Induction: Recursion

- Assume leaf nodes corresponding to information set at stage $t+1$ for $X$ have cost representation $\max_{\nu \in Q(I(X,t+1))} \langle \beta', \nu \rangle$

- Assume top level probabilities $\beta$ on $I^{X,t}$

- **Theorem:** The optimal value of the min-max game rooted at information set $I(X,t)$ with probabilities $\beta$ can be expressed as $\max_{\nu \in Q(I(X,t))} \langle \beta, \nu \rangle$
  - $Q(I(X,t))$ obtained from $Q(I(X,t+1))$ for information sets $I(X,t+1)$ that descend from $I(X,t)$
  - Key argument: convexifying space of support hyperplanes, embedding into a larger min-max game – treat the choice of hyperplane as a decision
Main Algorithm

- Recursive, backward induction procedure for obtaining values of subgames
  - Extension of POMDP algorithms to nested information games
  - Algorithm for computing support hyperplanes of cost-to-go function
- Can now solve for top level problem, get stage 1 saddle point strategies
  - Requires recursive solution of lower level problems as one-stage games \( \rightarrow \) Not as easy as dynamic programming
- **Theorem**: Behavior strategies obtained by the forward recursive solution are saddle point strategies in the original game
Example

- **Value estimation phase**

<table>
<thead>
<tr>
<th>$Q^j$</th>
<th>$W^j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G^{21}$</td>
<td>(1, 0)</td>
</tr>
<tr>
<td></td>
<td>(0, 2)</td>
</tr>
<tr>
<td>$G^{22}$</td>
<td>(2, 0)</td>
</tr>
<tr>
<td></td>
<td>(0, 1)</td>
</tr>
</tbody>
</table>

- **Saddle point retrieval phase**

\[\begin{align*}
\nu^1* &= (2/3, 1/3; 1/3, 2/3), \\
\omega^1* &= (2/3, 1/3; 1/3, 2/3; 0) \Rightarrow u^1* = (1, 0), \lambda^{*21}, \lambda^{*22}. \\
\beta^{21} &= (2/3, 1/3), \beta^{21} = (1/3, 2/3) \\
\omega^{21*} &= W^{21} \ast (\lambda^{*21}_1, \lambda^{*21}_2) = (0, 2/3; 0, 1/3) \Rightarrow u^{21*} = (2/3; 1/3); \\
\omega^{22*} &= W^{22} \ast (\lambda^{*22}_1, \lambda^{*22}_2) = (0, 1/3; 0, 2/3) \Rightarrow u^{22*} = (1/3; 2/3)
\end{align*}\]
Result

- Recursive, backward induction procedure for solving subgames
  - Extension of POMDP algorithms to nested information zero-sum games
  - Have algorithm for computing support hyperplanes of cost-to-go function
- Extended standard POMDP algorithms such as point-based value iteration to solution of games
- Difficulty
  - Still need to solve a value function for every information set of the uninformed player
  - Grows exponentially with number of stages...
Markov Structure

- Assume state $s^t$ at stage $t$, with Markov evolution depending on decisions by $X$ and $Y$, corresponding to Nature’s strategy:
  $$p(s^t | s^{t-1}, a^t, d^t)$$

- Assume cost structure additive and compatible with Markov structure
  $$J = \sum_{t=1}^{T-1} c^t(s^{t-1}, a^t, d^t) + c^T(s^T)$$

- Nested information structure
  - $Y$ has perfect information: at stage $t$, knows $(s^0, a^1, d^1, s^1, \ldots, s^{t-1}, a^t)$
  - $X$ observes only actions: at stage $t$, knows $(a^1, d^1, a^2, \ldots, d^{t-1})$
For every two histories in the same attacker information set $I_{X,t}$ at stage $t$, with same current state $\Rightarrow$ future costs and state evolution are the same given the same future actions

- The value of Markov states and costs
- Can aggregate same states with different histories in information set
**Theorem**: At any decision stage $t$, the solution of the subgame starting at $I_{X,t}^t$ is a piecewise linear, convex function of the distribution on internal states $\beta^t$.

**Theorem**: Consider two different information sets $I_{1}^{X,t}$ and $I_{2}^{X,t}$ of $X$ at the same stage. Then, the solution of their subgames have the same value function, which depends only on the distributions $\beta_{1,t}^1$, $\beta_{2,t}^2$.

- Only one value function per stage needed; no exponential growth (in value functions...not in number of hyperplanes!)
Experiments

- Multiple attacks on a simple network, one arc at a time
- Compare LP approach with nested decomposition technique

<table>
<thead>
<tr>
<th>Stages</th>
<th>Feasible attacks</th>
<th>LP run time (CPLEX)</th>
<th>Decom. run time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>150.9</td>
<td>0.22</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0.09</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1.83</td>
<td>0.30</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>&gt; 24 hours</td>
<td>0.83</td>
</tr>
</tbody>
</table>
Conclusions

- New approach for recursive solution of zero-sum dynamic games with nested information patterns
  - Extensions of POMDP algorithms
  - Novel approach for back substitution to get strategies
- Generalized theory and algorithms to Markov zero-sum games with nested information patterns
  - Much more efficient solution techniques
- Some extensions completed
  - Partial information on both players, still nested, and different orders
- Hard extensions
  - Non-nested information, many players, ...
  - Still hard to do for larger systems!
Thank you.

Congratulations, Demos!
Example: Cyber Attack

- Eavesdropper monitors traffic on subset of arcs
- Network using monitored arc confirms its presence; avoiding use of monitored arc costly
- Eavesdropper can change monitoring each period
- After K periods of monitoring, eavesdropper chooses where to attack for maximal degradation
- Payoff to network: Residual performance after attack minus cost of avoiding monitors

Numbers indicate capacity and probability arc is actually there