Stochastic Game for Deception and Self-Secured Cyber Physical Systems

Charles Kamhoua, PhD
Electronic Engineer
Army Research Laboratory, Network Security Branch

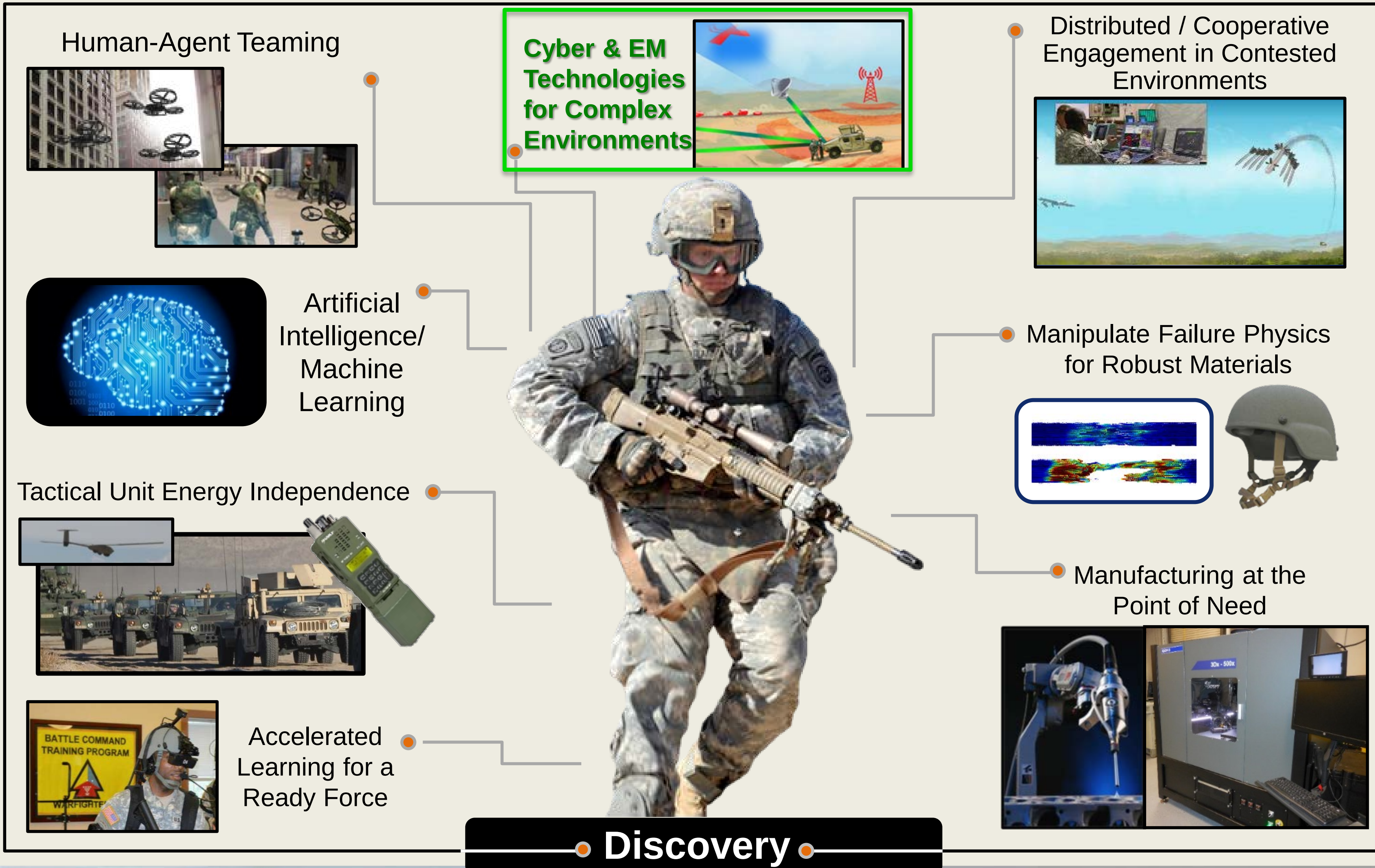
Definition

Cyber-Physical Systems or "smart" systems are co-engineered interacting networks of physical and computational components

Game Theory is the study of mathematical models of **conflict** and **cooperation** between **intelligent rational** decision-makers

U.S. ARMY RDECOM ARL

Essential Research Areas



The Nation's Premier Laboratory for Land Forces

U.S. ARMY RDECOM

CETCE ERA
Camouflage and Decoy of CEMA

ARL

CSA Priority: Network/C3I

Army Gaps

Maintaining a Robust Tactical Network: Tactical networked communication must persist in a cyber and EM contested and congested environments.

Defeating adversarial EM detection and kinetic targeting of communication systems: Future emitters must be able to communicate without reduced capability while being protected from adversary geolocation.

Protecting cyber-physical systems from targeted CEMA attacks: Electronics for combat vehicles and dismounts must be protected from CEMA attacks, using honeynets, decoys and camouflage.

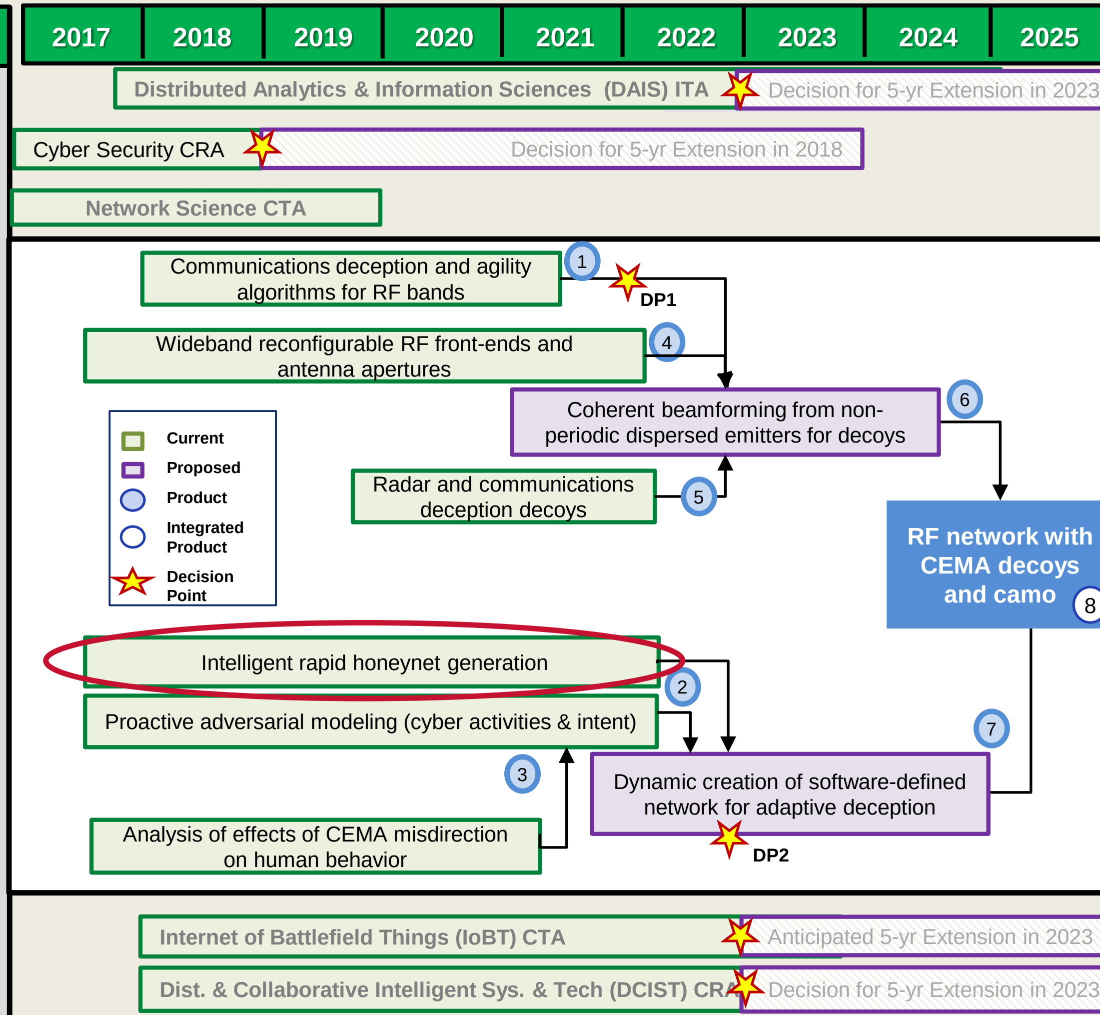
Outcomes and Products

Capabilities of CEMA Camo/Decoy

1. Network connectivity persists in the presence of attacks that disable up to three communication modalities
2. Maintain communication/ network range with 90% reduction in perceived RF signature, reducing adversarial geolocation capability for precision fires
3. Reduce operational effects from CEMA attacks to increase mission completion by 50% as compared to baseline

Products:

1. CEMA-camouflaged communications demonstration
2. Proactive network defense and resilience algorithms
3. Quantitative characterization and models of CEMA effects and adversarial intent
4. Situation-adaptive, multi-waveform, multi-function RF front-end demonstrator
5. Rapidly deployable radar/comms decoy with reduced SWaP
6. EM decoys that generate RF ghost images
7. Dynamic honeypot generation based on adversarial actions
8. RF network integrated with CEMA camo and virtual decoys



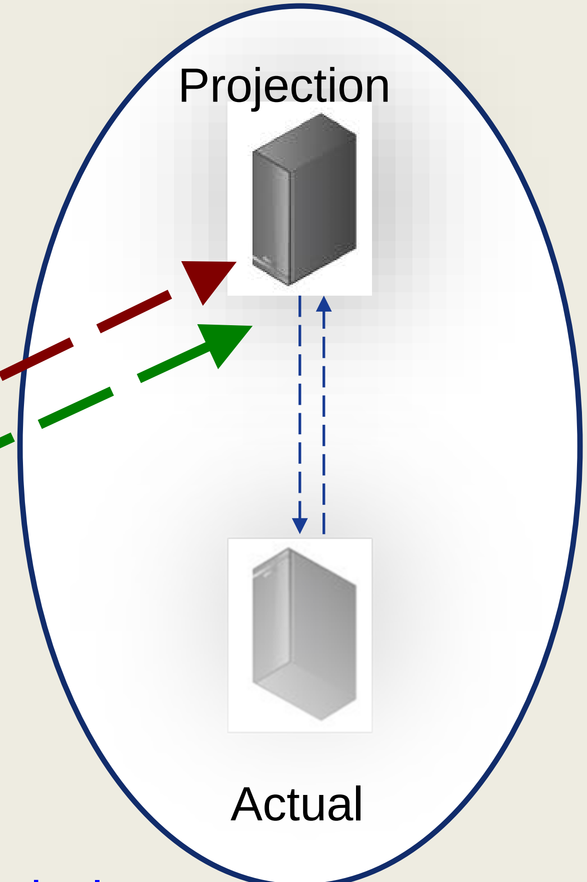
DP1: Reduction of 60% in perceived RF signature for communication
DP2: Deceive adversary to achieve 25% increase in mission completion compared to baseline

The Nation's Premier Laboratory for Land Forces

Cyber-CAMO System Overview

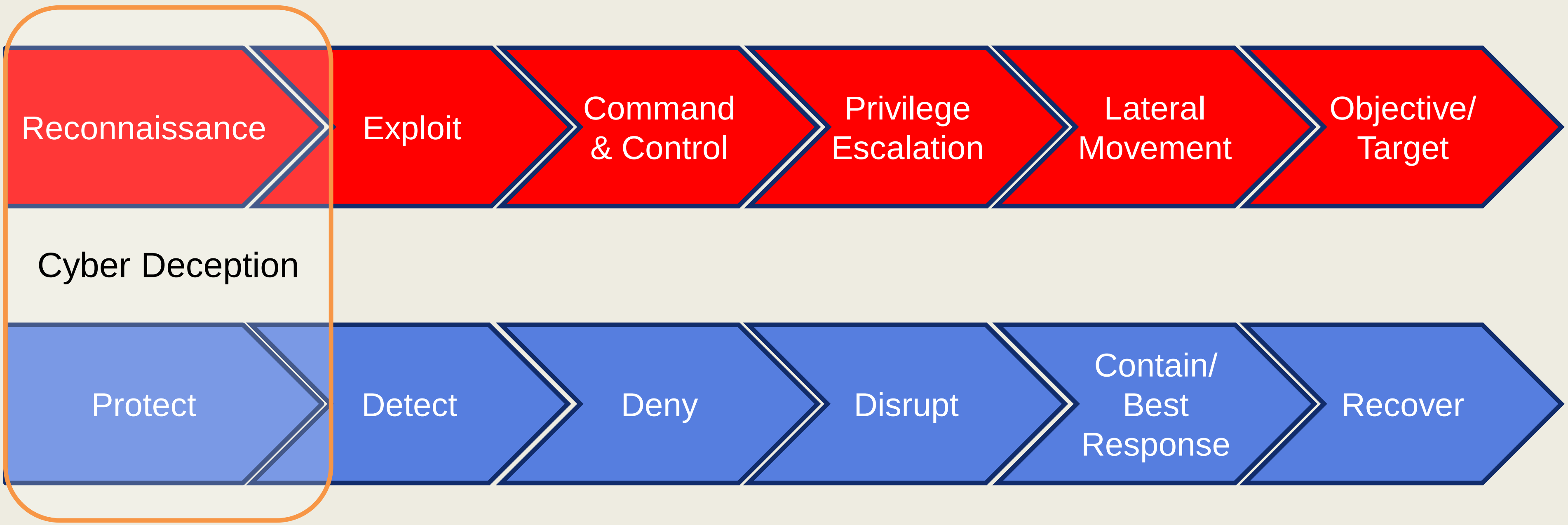
From afar, Adversary observes:

- Physical Camouflage: Actual target is projected onto one or more different geographic locations (***E/W CAMO***)
  - SEDD focus area
- Logical Camouflage: Actual cyber network component is dynamically projected onto one or more “honey-nets” (***Cyber CAMO***)
  - CISD/NSB focus area



- Final implementation may be combination of physical and logical camouflage

Cyber Kill Chain



Goal: Develop novel approaches to intelligently disguise a CPS network and impair the attacker's decision with false information to protect critical nodes.

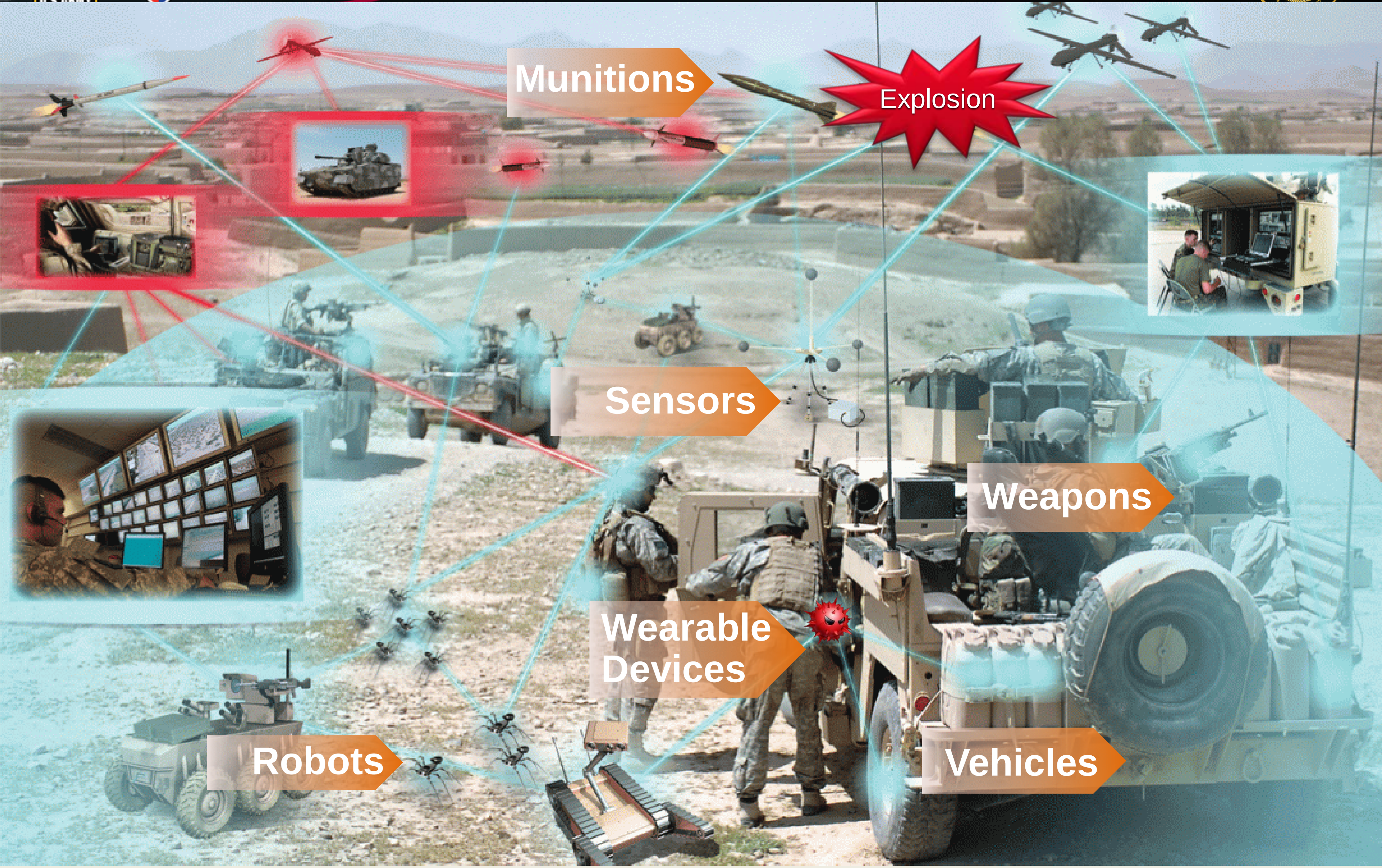
Research Challenge

- Limited battery power
- Limited computational power
- Low cost commercial off-the-shelf (COST) devise
- Heterogeneous devise designed with no security consideration
- Node mobility
- Contested and congested environment

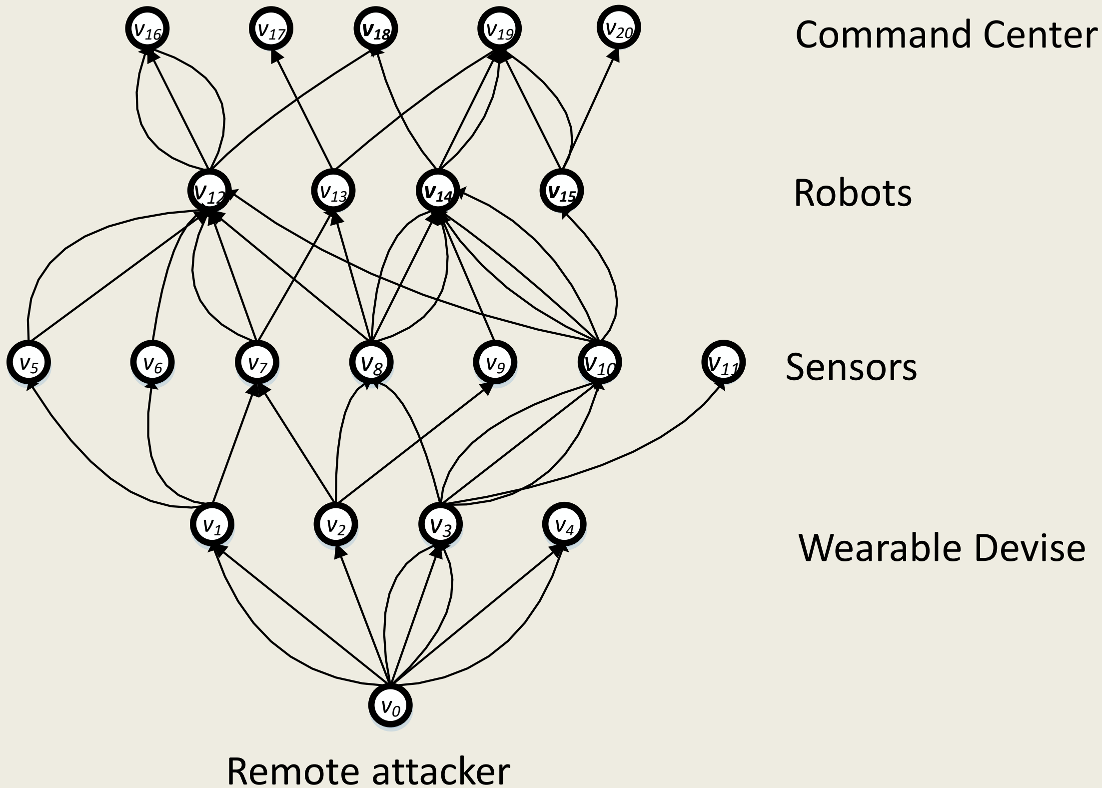
Related Works

- ❑ Updates system configuration based on risk [Zhu & Basar 2013]
- ❑ Consider the cost of mixed strategy in MTD [Rass et al. 2017]
- ❑ Deceptive routing against jamming attacks [Clark et al. 2012]
- ❑ Signaling game to disguise honeypots [Carroll and Grosu 2011]
- ❑ Bayesian honeypot selection by value [Kiekintveld et al. 2015]
- ❑ Signaling game for honeypot deployment [Pawlick & Zhu 2015]
- ❑ Stackelberg & attack graphs for deception [Durkota et al. 2015]
- ❑ Respond to attacker lateral movement [Mouhammad et al. 2016]

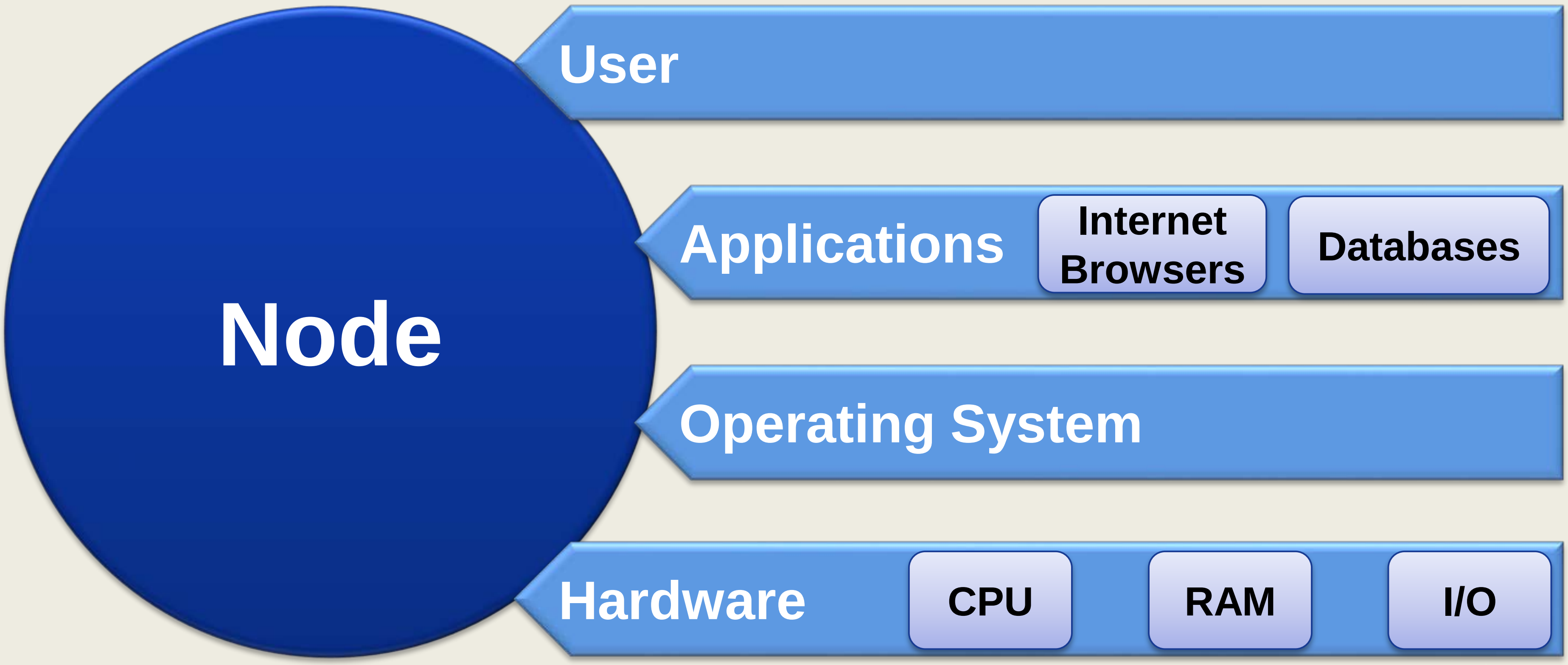
Scenario


Munitions
Explosion
Sensors
Weapons
Wearable Devices
Robots
Vehicles

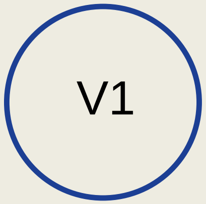
Attack Graph



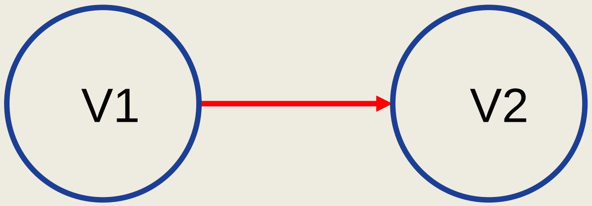
Node Composition



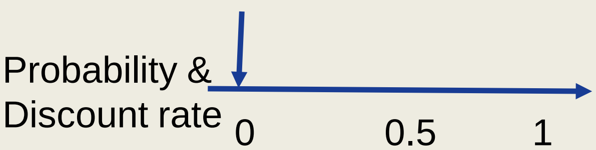
Type of Deception



- ❑ Add a fake **node**
- ❑ Hide critical **node**
- ❑ Increase/decrease the **value of any node**



- ❑ Add a fake **link/vulnerability**
- ❑ Hide a **link/vulnerability**
- ❑ Increase/decrease the **cost of a vulnerability**



- ❑ Increase/decrease the **transition probability**
- ❑ Increase/decrease the **monitoring probability**
- ❑ Increase/decrease the **discount factor/rate**

Detected or Cover up?	Random or Deterministic?
Military or Civilian?	Rational or Irrational?

- ❑ **Deterministic vs random** network (MTD)
- ❑ Attacker **detected vs cover up**
- ❑ Hide **network identity**, e.g Military vs civilian
- ❑ **Full or limited rationality of users/software**
- ➢ computing power, memory space, data, algorithm

Vulnerability Multi-Graph

A ***vulnerability multi-graph*** $\boldsymbol{G(V, E)}$ *is a graph which depicts ways in which an adversary can exploit sequentially different vulnerabilities to break the system.* $V = \{v_1, \ldots, v_N\}$ *represents the set of nodes and* N *the total number of nodes.* $E \subseteq VxV$ *is the set of directed edges.*

- ❑ Each node v has a set of applications
- ❑ Each application has a set of known vulnerability (empty or not) and open ports through which illegitimate users may gain access to v
- ❑ Two nodes v_1 and v_2 are connected on G if it exists on node v_2 an application hosting a vulnerability that the system rules allow to access from node v_1.

Two-player Zero-sum Markov Game

A two-player zero sum Markov game is defined as a 6-tuple $(S, A, O, P, \mathcal{R}, \gamma)$ where:

- $S = \{s_1 .. s_l\}$ is a finite set of game states;
- $A = \{a_1 .. a_n\}$ is the set of actions of the maximizer (row player);
- $O = \{o_1 .. o_m\}$ is the set of actions of the minimizer (column player);
- P is a Markovian transition model, with $P(s, a, o, s')$ being the probability that s' will be the next game state when players take actions a and o respectively;
- The function $\mathcal{R}(s, a, o)$ specifies the immediate reward (or cost) of players for taking actions a and o in state s;
- $\gamma \in]0, 1]$ is the discount factor for future rewards.

Player's Policy

- A policy $\pi_A: S \rightarrow \Omega(A)$, for the row player (maximizer) is a function that gives for each state s a probability distribution $\pi_A(s)$ over the maximizer actions $A = \{a_1..a_n\}$. For any policy π_A, $\pi_A(s,a)$ denotes the probability to take action a in state s.

- For any policy π, $Q^\pi(s,a,o)$ is the expected sum of discounted reward of the row player:

$$Q^\pi(s,a,o) = \underbrace{\mathcal{R}(s,a,o)}_{\text{Immediate reward}} + \underbrace{\gamma \sum_{s' \in S} P(s,a,o,s') \min_{o' \in O} \sum_{a' \in A} Q^\pi(s,a,o)\, \pi(s',a')}_{\text{Future rewards}}$$

- Optimal policy:

$$\begin{cases} W(s) = \max\limits_{\pi_A(s) \in \Omega(A)} \min\limits_{o \in O} \sum_{a \in A} Q(s\ a,o)\, \pi'(s,a) \\ Q(s,a,o) = \sum_{s' \in S} P(s' |\, a,o,s)[\mathcal{R}(s,a,o,s') + \gamma W(s')] \end{cases}$$

Game Matrix

Reward matrix for state $s \in S$

		Column player			
		o_1	o_2	…	o_m
Row player	a_1	$Q(s, a_1, O_1)$			
	a_2				
	…				
	a_n				$Q(s, a_n, o_m)$

Value Iteration Algorithm

Value iteration $(S, A, O, P, \mathcal{R}, \gamma)$

$W \leftarrow 0$
$l \leftarrow 0$

Repeat

$l++$

For each $s \in S$ **do**

$$W_{l+1}(s) = \max_{\pi_A(s) \in \Omega(A)} \min_{o \in O} \sum_{a \in A} \pi(s,a) \sum_{s' \in S} P(s' | a, o, s)[\mathcal{R}(s,a,o,s') + \gamma W_l(s')]$$

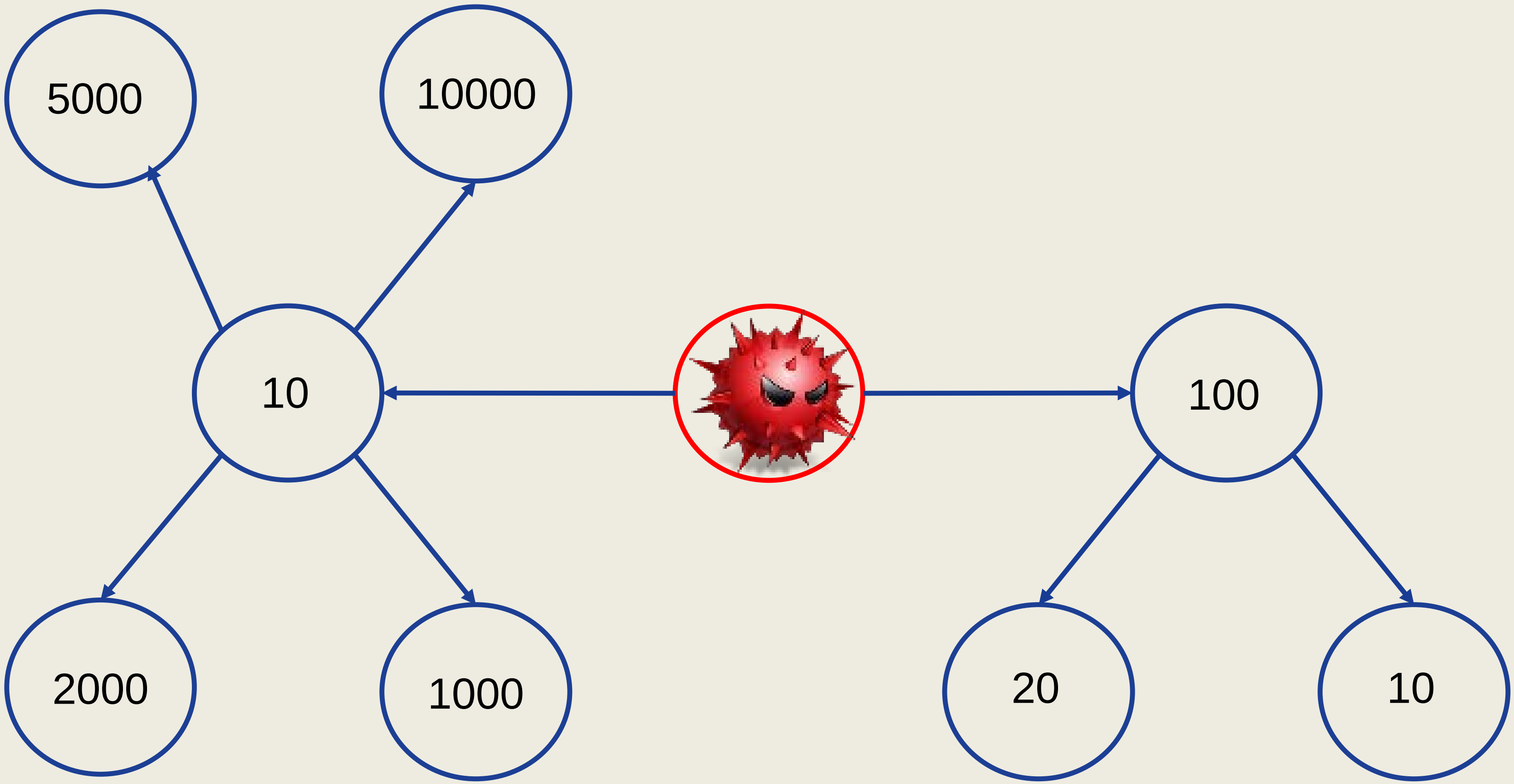
Until $\forall s \in S, |W_{l+1}(s) - W_l(s)| < \epsilon$

For each $s \in S$ **do**

$$\pi(s) \leftarrow \pi(s): \max_{\pi_A(s) \in \Omega(A)} \min_{o \in O} \sum_{a \in A} \pi(s,a) \sum_{s' \in S} P(s' | a, o, s)[\mathcal{R}(s,a,o,s') + \gamma W_l(s')]$$

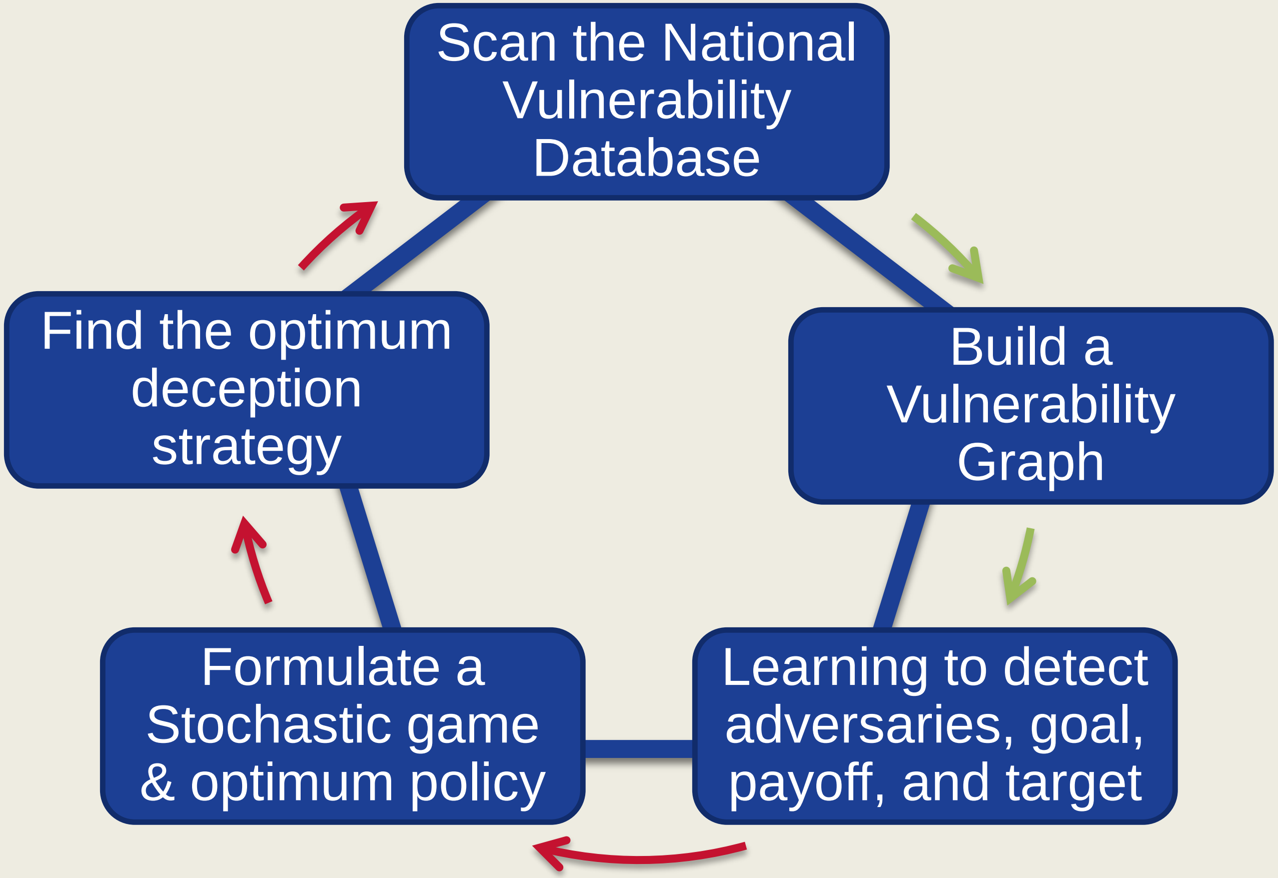
Return π, W_{l+1}

Identification of critical node for intelligent deception



What node is more attractive to the attacker? **Left or right?**

Game theory for Automated Deception



Measure of the Value of Cyber Deception

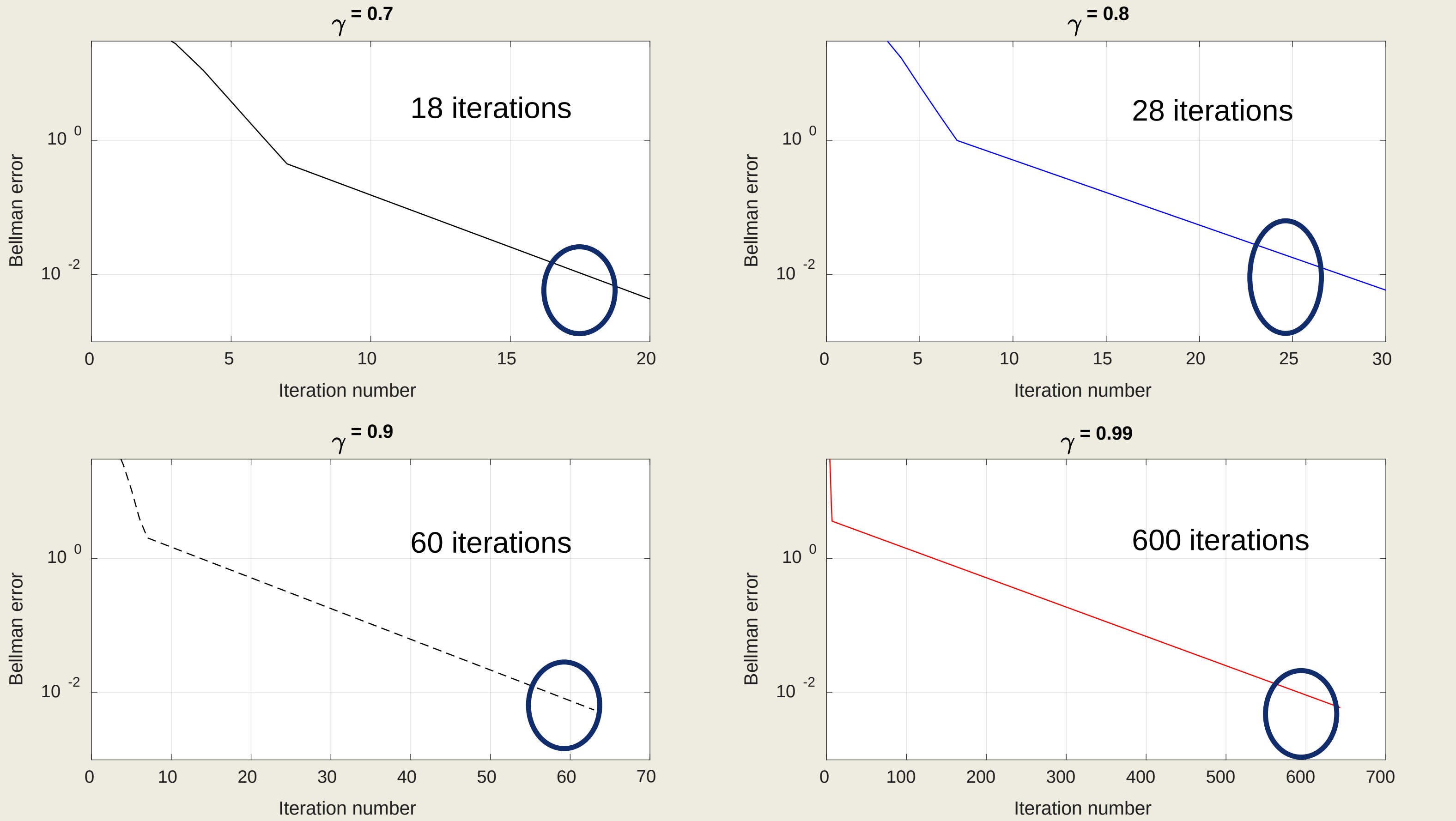
The value of cyber deception can be measured as the difference between:

The attacker's payoff in a game of complete information (No deception)

And

The attacker's payoff in that game after the defender apply cyber deception

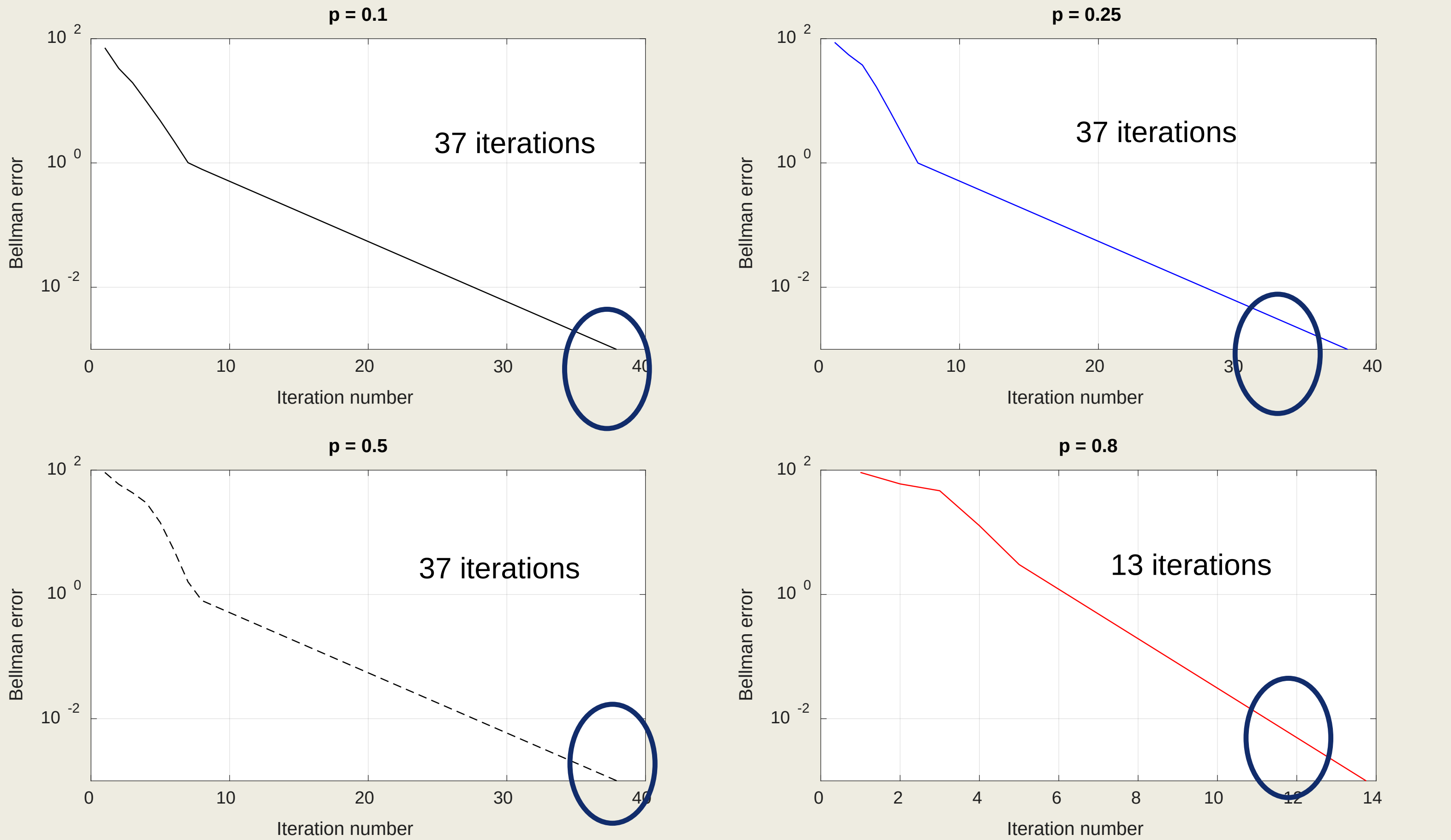
Convergence Speed vs Discounted Factor



The convergence speed is affected by the discounted factor.
The Bernoulli trial probability is $p = 0.4$ and the threshold error is 0.01

Validation

Convergence Speed vs Bernoulli Trial Probability



The convergence speed is less affected by the Bernoulli trial probability.
The discounted factor is is $\gamma = 0.8$ and the threshold error is 0.01

Deterministic Strategies

If the attacker uses a deterministic strategy, the optimal defense strategy is also deterministic and the attacker never succeed.

Attacker Strategy	**Optimal Defense Strategy**
Shortest path	Vulnerabilities corresponding to the shortest path
Least cost edges	Vulnerabilities corresponding to least cost edges
Movement toward next most attractive node	Vulnerabilities corresponding to most attractive node

The optimum policy is a mixed strategy at each state of the game

Future Works

- Imperfect monitoring
- Incomplete information
- Learning the attacker's attack graph
- Attacker's goal recognition
- Limited rationality
- Multiple colluding attacker
- Time varying attack graph
- Distributed defense mechanism

References

Quanyan Zhu and Tamer Basar, “Game-theoretic approach to feedback-driven multi-stage moving target defense”, in Decision and Game Theory for Security. Springer, 246–263, 2013.

Stefan Rass, Sandra Konig, Stefan Schauer, “On the Cost of Game Playing: How to Control the Expenses in Mixed Strategies”, in Decision and Game Theory for Security. Springer, 495–505, 2017.

Andrew Clark, Quanyan Zhu, Radha Poovendran, Tamer Basar, “Deceptive routing in relay networks”, in Decision and Game Theory for Security. Springer, 171–185, 2012.

Thomas E Carroll and Daniel Grosu, “A game theoretic investigation of deception in network security” Security and Commun. Nets. 4, 10, 1162–1172, (2011)

Christopher Kiekintveld, Viliam Lisy, and Radek Pibil, “Game-theoretic foundations for the strategic use of honeypots in network security” in Cyber Warfare, Springer, 81–101, 2015.

Jeffrey Pawlick and Quanyan Zhu, “A Mean-Field Stackelberg Game Approach for Obfuscation Adoption in Empirical Risk Minimization”, in Global Signal and Inform. Processing Workshop on Control and Game Theoretic Approaches to Security and Privacy, 2017.

A. Mouhammad, A. Fawaz, W. H. Sanders and T. Basar, "A Game-Theoretical Approach to Respond to Attacker Lateral Movement", International Conference on Decision and Game Theory, 2016.