Colonel Blotto Game Aided Attack-Defense Analysis in Real-World Networks

Sanghai Guan[1], Jingjing Wang[1], Chunxiao Jiang[1],
Zhu Han[2], Yong Ren[1], and Abderrahim Benslimane[3]

[1]Tsinghua University, Beijing, China
[2]University of Houston, Houston, TX, USA
[3]University of Avignon, Avignon, France

gsh17@mails.tsinghua.edu.cn
jchx@tsinghua.edu.cn, zhan2@uh.edu, reny@tsinghua.edu.cn,
abderrahim.benslimane@univ-avignon.fr

IEEE GLOBECOM, December 12, 2018

Content

Content

Introduction

Motivation:

- **Network systems**, such as Internet, smart grids, transportation networks, social networks, etc., play a critical role in human society.
- However, due to their **inherent vulnerability** as well as the limited management and operational capability, these network systems are constantly under the threat of malicious attackers.
- Therefore, in such attack-defense scenarios, it is particularly significant to give **precise analysis** and make the best use of **limited resources**.

Introduction

Attack-defence resource allocation & Colonel Blotto game:

- **Colonel Blotto game** is a useful model for attack-defense resource allocation, where two players have to allocate limited troops on several battlefields.
- In Colonel Blotto game, a player wins a battlefield if he assigns more troops on it than his counterpart. The goal of both players is to win as many battlefields as possible.
- It has been widely studied and applied in **a range of fields** such as military, information forecasting, social science, communication and computer networks, etc.

Introduction

Challenges:

- Existed models just establish a **simple and linear relationship** between the global utility and the results on each battlefield. In practical networks systems, the global utility and the result of each battlefield often have a **complex and implicit relationship**.
- With the increase of the number of troops and battlefields, the number of feasible actions grows exponentially. Hence, most related works just concentrate on simple toy systems. **Efficient solutions for large scale network systems** are urgently needed.

Introduction

Our original contributions:

- **Networked Colonel Blotto game model** for **attack-defense resource allocation** in network systems, including four metrics that evaluate network performance and formulate the utility of this **two-player zero-sum game**.
- A genetic algorithm based **co-evolution algorithm** for searching quality strategies for both players which reduces the complexity of finding the equilibrium.
- Applying our proposed game model to four large-scale network systems, i.e., **Internet**, **vehicular networks**, **air transportation systems** and **social networks**, in simulation.

Content

Game Model

The **networked Colonel Blotto game** is a **one-shot two-player zero-sum game**, where two players are the **defender** and the **attacker**, respectively.

Network Model:

$\boldsymbol{G} = \{\mathbb{V}, \mathbb{E}\}$: The network system defined as an undirected graph.

$\mathbb{V} = \{v_1, v_2, \ldots, v_N\}$: The set of nodes.

N: The total number of nodes.

$\mathbb{E} = \{e_1, e_2, \ldots, e_M\}$: The set of edges.

M: The total number of edges.

$e_k = \{v_i, v_j\}$: The edge that connects nodes v_i and v_j.

Game Model

Resource Allocation:

A_1: The quantity of defense resources for the defender.

A_2: The quantity of defense resources for the attacker.

$\boldsymbol{a}_1 = [a_1^1, a_1^2, \ldots, a_1^N]$: The action of the defender.

$\boldsymbol{a}_2 = [a_2^1, a_2^2, \ldots, a_2^N]$: The action of the attacker.

$a_l^i \geq 0$ $(l = 1, 2)$ stands for the quantity of resources allocated on node v_i by players and $\sum_{i=1}^N a_l^i = A_l$ $(l = 1, 2)$.

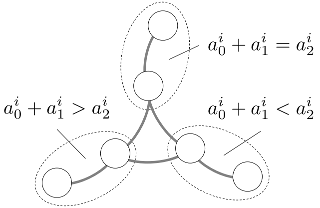
$\boldsymbol{a}_0 = [a_0^1, a_0^2, \ldots, a_0^N]$: Nodes' self-defense capability $(a_0^i \geq 0)$.

Game Model

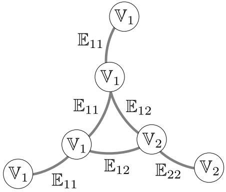
Game Rule:

The result of the "battle" on each node depends on the quantity of the attack-defense resources that two players allocate.

- Set of nodes $\mathbb{V}_1$ and $\mathbb{V}_2$.
- Set of edges $\mathbb{E}_{11}$, $\mathbb{E}_{12}$ and $\mathbb{E}_{22}$.



(a) Resources allocated



(b) Nodes' and edges' affiliation

Figure: The relationship between the nodes' attack-defense resources allocated and their affiliation.

Game Model

Utility Function:

In order to compare the performance of the whole network system, we denote the **original network as G'**, while the **network after the game is G''**. The utility function can be given by:

$$u_1(\boldsymbol{a}_1, \boldsymbol{a}_2) = -u_2(\boldsymbol{a}_1, \boldsymbol{a}_2) = f(\boldsymbol{G}'') - f(\boldsymbol{G}'),$$

u_1, u_2: the utility of the defender and the attacker.

$f(\cdot)$: the evaluation function of the network performance.

Game Model

Utility Function:

In order to compare the performance of the whole network system, we denote the **original network as G'**, while the **network after the game is G''**. The utility function can be given by:

$$u_1(\boldsymbol{a}_1, \boldsymbol{a}_2) = -u_2(\boldsymbol{a}_1, \boldsymbol{a}_2) = f(\boldsymbol{G}'') - f(\boldsymbol{G}'),$$

u_1, u_2: the utility of the defender and the attacker.
$f(\cdot)$: the evaluation function of the network performance.

In original network system $\boldsymbol{G}'$, we assume that $a_1^i = a_2^i = 0$, so all the nodes in $\boldsymbol{G}'$ belong to $\mathbb{V}_1$. Therefore, the **defender's goal** is to **minimize the performance loss**, while the **attacker** aims for **maximizing it**, which constitutes a **zero-sum game**.

Network Performance Metrics

For the convenience of deduction, we adopt the **adjacency matrix** as $\boldsymbol{W} = (w_{ij})_{N \times N}$ to represent the network topology, i.e.,

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{bmatrix}.$$

- In an unweighted graph, $w_{ij} \in \{0, 1\}$ represents the existence of edge $\{v_i, v_j\}$.
- In a weighted graph, $w_{ij} \geq 0$ denotes the weight of edge $\{v_i, v_j\}$.

Network Performance Metric I: Network Connectivity

- If some nodes are controlled and damaged by the attacker, the **network connectivity** will seriously change.
- The survivability of the network system, i.e., the ability of maintaining its connectivity, becomes an critical metric.

For an unweighted graph, the weight of edge w_{ij} can be defined as:

$$w_{ij} = \begin{cases} 1, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{11}, \\ 0, & \text{if } \{v_i, v_j\} \notin \mathbb{E}_{11}. \end{cases}$$

The network can be divided into one or more sub-networks. The sub-network with most nodes is called the **giant component**.

Network Performance Metric I: Network Connectivity

- If some nodes are controlled and damaged by the attacker, the **network connectivity** will seriously change.
- The survivability of the network system, i.e., the ability of maintaining its connectivity, becomes an critical metric.

For an unweighted graph, the weight of edge w_{ij} can be defined as:

$$w_{ij} = \begin{cases} 1, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{11}, \\ 0, & \text{if } \{v_i, v_j\} \notin \mathbb{E}_{11}. \end{cases}$$

The network can be divided into one or more sub-networks. The sub-network with most nodes is called the **giant component**.

If the giant component contains n nodes, the network connectivity based evaluation function can be denoted as: $f(\boldsymbol{G}) = n$.

Network Performance Metric II: Average Path Length

- Sometimes, attacks may not damage the network's connectivity, but may still influence **performance of edges**.

$\boldsymbol{p}_{i_1,i_K} = [v_{i_1}, v_{i_2}, \ldots, v_{i_K}]$: Path between nodes v_{i_1} and v_{i_K}.

$r(\boldsymbol{p}_{i_1,i_K}) = \sum_{[v_{i_k}, v_{i_{k+1}}] \in \boldsymbol{p}_{i_1,i_K}} w_{i_k,i_{k+1}}$: The length of a path.

$r^\star_{ij} = \min_{\boldsymbol{p}_{ij}} r(\boldsymbol{p}_{ij})$: The shortest path length between two nodes.

$\bar{r} = \frac{\sum_{i \neq j} r^\star_{ij}}{N(N-1)}$ The average path length of the network.

Network Performance Metric II: Average Path Length

- Sometimes, attacks may not damage the network's connectivity, but may still influence **performance of edges**.

$\boldsymbol{p}_{i_1,i_K} = [v_{i_1}, v_{i_2}, \dots, v_{i_K}]$: Path between nodes v_{i_1} and v_{i_K}.

$r(\boldsymbol{p}_{i_1,i_K}) = \sum_{[v_{i_k}, v_{i_{k+1}}] \in \boldsymbol{p}_{i_1,i_K}} w_{i_k,i_{k+1}}$: The length of a path.

$r^\star_{ij} = \min_{\boldsymbol{p}_{ij}} r(\boldsymbol{p}_{ij})$: The shortest path length between two nodes.

$\bar{r} = \frac{\sum_{i \neq j} r^\star_{ij}}{N(N-1)}$ The average path length of the network.

The average path length based evaluation function can be formulated as: $f(\boldsymbol{G}) = -\bar{r}$.

Network Performance Metric III: Average Degree

- **Degree** is a critical and universal metric of a network system which reveals its connectivity, structure, or other characteristics.

$d_i = \sum_{j=1}^{N} w_{ij}$: The degree of node v_i.

$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}}{N}$: The average degree of a network.

Network Performance Metric III: Average Degree

- **Degree** is a critical and universal metric of a network system which reveals its connectivity, structure, or other characteristics.

$d_i = \sum_{j=1}^{N} w_{ij}$: The degree of node v_i.

$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}}{N}$: The average degree of a network.

The average degree based evaluation function of the network system can be defined as: $f(\boldsymbol{G}) = \bar{d}$.

Network Performance Metric IV: Transmission Capability

- Some transmission processes, such as rumors in social networks, disease in the crowd and computer virus in computer networks, can be harmful.
- The **susceptible-infection (SI) propagation model** is commonly used.

t The time step of transmission process. We take the time after the game as time step $t = 0$.

$\mathbb{V}_1(t)$, $\mathbb{V}_2(t)$ The susceptible node set and the infected node set at time step t.

$\mathbb{E}_{ij}(t)$ The edge sets at time step t.

Network Performance Metric IV: Transmission Capability

Game rule:

- Nodes controlled by the defender constitute $\mathbb{V}_1(0)$.
- Nodes controlled by the attacker constitute $\mathbb{V}_2(0)$.

At each time step t, node v_i may be infected and added into $\mathbb{V}_2(t)$ with the probability of:

$$p_i(t) = \begin{cases} \dfrac{\sum_{\{j:\ \{v_i,v_j\}\in\mathbb{E}_{12}(t-1)\}} c_j}{\sum_{\{j:\ \{v_i,v_j\}\in\mathbb{E}\}} c_j}, & \text{if } v_i \in \mathbb{V}_1(t-1), \\ 1, & \text{if } v_i \in \mathbb{V}_2(t-1), \end{cases}$$

where c_j is defined as the **influence** of node v_j.
Correspondingly, node v_i may stay susceptible and fall into $\mathbb{V}_1(t)$ with probability $1-p_i(t)$.

Network Performance Metric IV: Transmission Capability

We define the **average diffusion time** $\bar{t}$ as the expected time when the proportion of infected nodes reaching a threshold β, i.e.,

$$\bar{t} = \mathbf{E}\left(\min\left\{t\colon \frac{|\mathbb{V}_2(t)|}{N} \geq \beta\right\}\right).$$

where $|\mathbb{V}_2(t)|$ represents the number of infected nodes at t.

Network Performance Metric IV: Transmission Capability

We define the **average diffusion time** $\bar{t}$ as the expected time when the proportion of infected nodes reaching a threshold β, i.e.,

$$\bar{t} = \mathbf{E}\left(\min\left\{t \colon \frac{|\mathbb{V}_2(t)|}{N} \geq \beta\right\}\right).$$

where $|\mathbb{V}_2(t)|$ represents the number of infected nodes at t.

The transmission capability based performance evaluation function can be given by: $f(\boldsymbol{G}) = \bar{t}$.

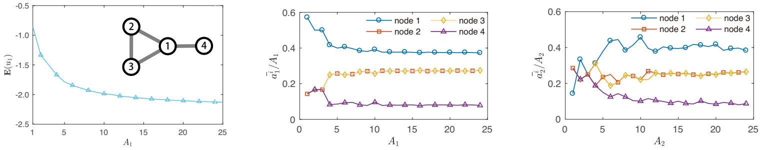
Content

Strategies in Small-Scale Network Systems

- The proposed game is with **infinite actions and discontinuous payoff**, which brings difficulties to the analysis.
- Use **gridding method** to transform it into the game with finite actions. If its equilibrium is insensitive to different gridding, it will approximate the original equilibrium gradually with finer and finer grid density.
- When the grid density approaches to infinity, it will **converge to the equilibrium** of the original game.

Gridding Based Equilibrium Solution Algorithm

The transformed zero-sum game with finite actions can be solved by linear programming.



(a) Expected Utility (b) Defender's strategy (c) Attacker's strategy

Figure: The **convergence** of expected utility and normalized average resource allocation on each node under approximated mixed Nash equilibrium strategy with **finer grid density**.

However, its computational complexity raises rapidly with the increase of network scale (factorial with N) .

Strategies in Large-Scale Network Systems

- In real network systems, attackers and defenders also have several **commonly used patterns** for attacking and defending. These specific patterns can be regarded as the common chosen actions in the experiments.
- The rational defender and attacker will only choose the **actions yielding high expected utility** as its strategy.
- Therefore, in order to simplify the computation, we assume that the action set of the player is composed of only a small part of the quality practical actions from all the feasible actions, namely **the practical action set**.

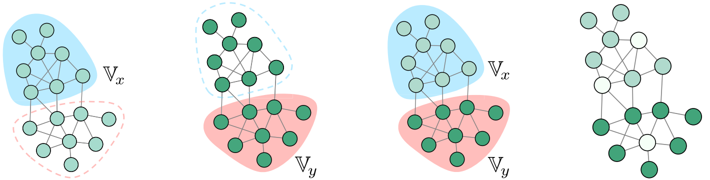
Co-Evolutionary Based Algorithm for Large-Scale Network Systems

The process of co-evolution:

- Generate some random actions constituting the initial action set for the defender and the attacker.
- **Genes**: $\boldsymbol{g}_l = [g_l^{(1)}, g_l^{(2)}, \ldots, g_l^{(N)}]$, $g_l^{(i)} \geq 0$ $(i = 1, 2, \ldots, N)$ are random numbers.
- **Actions**: $\boldsymbol{a}_l^k = A_l \cdot \frac{\boldsymbol{g}_l^k}{\sum_{i=1}^N g_l^{k(i)}}$ (For the attacker, The resources on node that $a_2^i \leq a_0^i$ will be re-allocated).
- The defender and the attacker test these actions by matching against the opponents' action sets and record the average utility of each action.

Co-Evolutionary Based Algorithm for Large-Scale Network Systems

- The actions with high average utility will be added directly into the next generation, and the other actions in the next generation will be generated by genetic manipulation, i.e., **crossover** and **random mutation**.



(a) parent chromosome (b) parent chromosome (c) crossover chromosome (d) child chromosome

Figure: The process of generating child chromosome from parent chromosomes.

Co-Evolutionary Based Algorithm for Large-Scale Network Systems

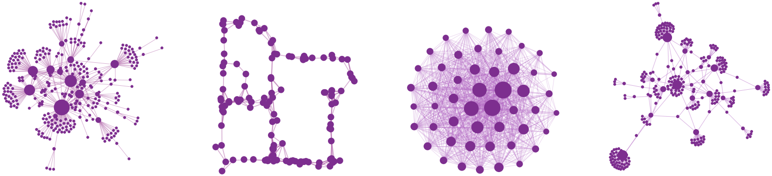
The process of co-evolution:

- In such an iterative process, dominated actions will be continuously excluded from the action set, and quality actions can still be retained.
- Actions with higher quality can be generated by genetic manipulation, which yields the co-evolution of both players' action sets.
- Finally, we can take the result as the practical action sets for both players and solve the equilibrium.

Content

Applications and Simulations

- Selected applications of our game model in realistic scenarios based on **real-world large-scale networks**.
- Simulations based on the **co-evolution based algorithm**.
- **Four scenarios** correspond to **four performance metrics**.



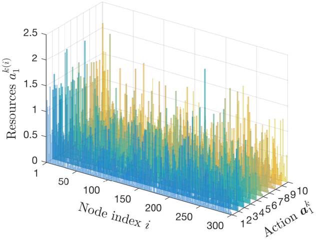
(a) Internet (b) Vehicle network (c) Air network (d) Social network

Figure: Real-world networks for simulation.

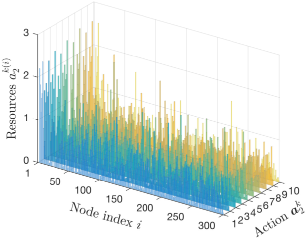
Simulation I: Internet Security

- **Attackers** can attack key network devices in Internet by distributed denial of service (DDoS), identity spoofing, intrusion, etc.
- **Defenders** can protect network devices by installing firewalls, upgrading hardwares and softwares, and so on.
- **Network Data**: University of Oregon Route Views Project, 300 nodes (Internet autonomous systems) , 400 edges (network routes), undirected and unweighted network.
- **Network performance metric**: Network Connectivity $f(\boldsymbol{G}) = n$.
- **Game rule**: $w_{ij} = \begin{cases} 1, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{11}, \\ 0, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{12} \cup \mathbb{E}_{22}. \end{cases}$

Simulation I: Internet Security



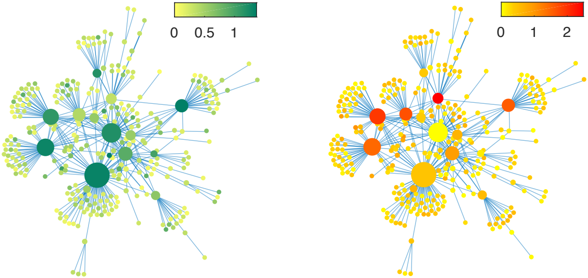
(a) Defender



(b) Attacker

Figure: The practical action sets of the defender and the attacker. derived from co-evolution algorithm when $A_1 = A_2 = 100$ and $a_0^i = 0.01 \cdot d_i$ (The nodes' indices are sorted by degree in descending order).

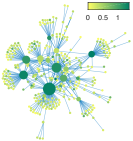
Simulation I: Internet Security



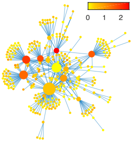
(a) Defender (b) Attacker

Figure: Expected resource allocation of the defenders and attackers when $A_1 = A_2 = 100$ and $a_0^i = 0.01 \cdot d_i$.

Simulation I: Internet Security



(a) Defender



(b) Attacker

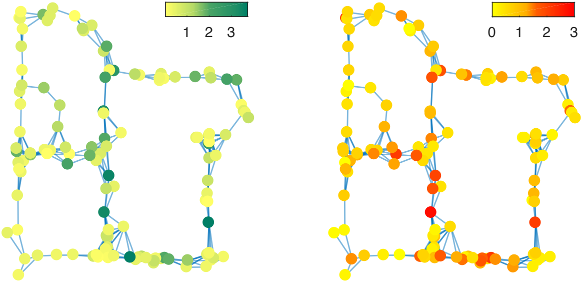
Figure: Result

- $\mathbf{E}(u_1) = -198.5$ when $A_1 = A_2 = 100$.
- The attacker tends to allocate much resources on nodes with **high degree**, which makes their neighboring nodes separated from the giant component, and on nodes with **high centrality** to make the whole network collapse.
- Because there are a few nodes with large degree and there exist hierarchical structures, this network is **vulnerable** to targeted attacks.

Simulation II: Communication Timeliness of Vehicular Networks

- **Attackers** can interfere the communication of some vehicle devices through jamming.
- **Defenders** (staff or softwares) can increase the transmission power and improve anti-interference capacity of these devices.
- **Network Data**: Beijing Taxi GPS Dataset in T-Drive Project, 125 nodes (taxis) , 420 edges (wireless connections), undirected and unweighted network.
- **Network performance metric**: Average path length $f(\boldsymbol{G}) = -\bar{r}$.
- **Game rule**: $w_{ij} = \begin{cases} 1, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{11}, \\ 10, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{12} \cup \mathbb{E}_{22}. \end{cases}$

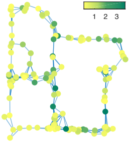
Simulation II: Communication Timeliness of Vehicular Networks



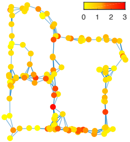
(a) Defender (b) Attacker

Figure: Expected resource allocation of the defenders and attackers when $A_1 = A_2 = 100$ and $a_0^i = 0.1$.

Simulation II: Communication Timeliness of Vehicular Networks



(a) Defender



(b) Attacker

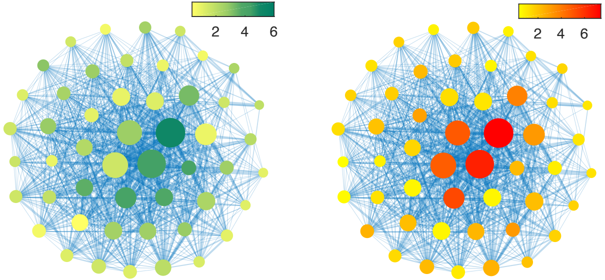
Figure: Result

- Both players tend to allocate more resources on the nodes with **high centrality**.
- The **gateway nodes**, which are the nodes must be passed in numerous shortest paths, play an important role
- Increasing the **density of vehicles** or **vehicles' maximum communication distance** will create more links between vehicles, which is beneficial for improving the anti-interfere capacity and timeliness of communication.

Simulation III: Efficiency and Reliability of Transportation Systems

- **Attackers** can obstruct airline schedules by causing terrorist attacks, accidents and havoc.
- **Defenders** can improve airports' prevention and response capacity to various risks.
- **Network Data**: US Air Transportation Network Dataset, 50 nodes (airports) , 878 edges (flights), undirected and weighted network.
- **Network performance metric**: Average degree $f(\boldsymbol{G}) = \bar{d}$.
- **Game rule**: $w''_{ij} = \begin{cases} w'_{ij}, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{11}, \\ \frac{1}{2} w'_{ij}, & \text{if } \{v_i, v_j\} \in \mathbb{E}_{12} \cup \mathbb{E}_{22}. \end{cases}$

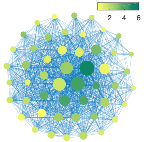
Simulation III: Efficiency and Reliability of Transportation Systems



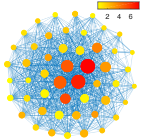
(a) Defender (b) Attacker

Figure: Expected resource allocation of the defenders and attackers when $A_1 = A_2 = 100$ and $a_0^i = 10^{-8} \cdot d_i$.

Simulation III: Efficiency and Reliability of Transportation Systems



(a) Defender



(b) Attacker

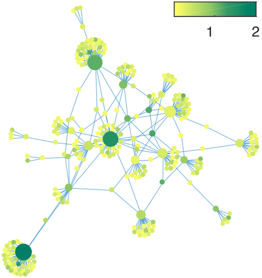
Figure: Result

- Both players tend to allocate more resources on the node with a **high degree**.
- Because this air network is **dense**, the results on it is similar to the case of traditional Colonel Blotto games with weighted battlefields.

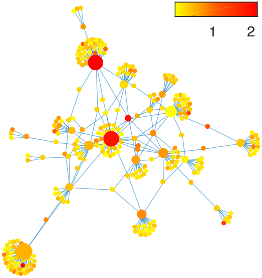
Simulation IV: Rumor Spread Control in Social Networks

- **Attackers** can spread rumors to users and turn them into initial rumor disseminators.
- **Defenders** can increase the resistance and discernment to rumors of social network users by opinion supervision.
- **Network Data**: Microblog PCU Dataset, 279 nodes (Weibo users) , 313 edges ("following each other" relationships), undirected and unweighted network.
- **Network performance metric**: Transmission capability $f(\boldsymbol{G}) = \bar{t}$.
- **Game rule**: SI model based diffusion, using the betweenness centrality to denote the influence c_k.

Simulation IV: Rumor Spread Control in Social Networks



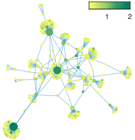
(a) Defender



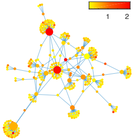
(b) Attacker

Figure: Expected resource allocation of the defenders and attackers when $A_1 = A_2 = 100$ and $a_0^i = 0.01 \cdot d_i$ (Assuming $f(\boldsymbol{G}') = 0$ for the convenience of elaborating).

Simulation IV: Rumor Spread Control in Social Networks



(a) Defender

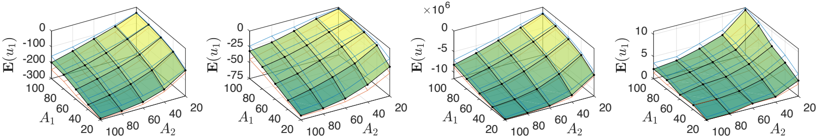


(b) Attacker

Figure: Result

- Both players mainly focus on two kinds of nodes. One is the nodes with **high influence**. The other is the **hub nodes** connecting the small sub-communities, which also play critical roles in rumor spread.
- Social network of friends has **strong transmission capacity**, and it is difficult for the defender to suppress the emergence and spread of rumors.

Summary



(a) Internet (b) Vehicle network (c) Air network (d) Social network

Figure: Expected utility of the defender under different A_1 and A_2.

The practical action set generated by the co-evolution algorithm overwhelms the randomly generated action set, which reveals the **effectiveness and validity** of our proposed algorithm.

Content

Conclusions

- Modeled the attack-defence resource allocation as a **networked zero-sum Colonel Blotto game**, which broadens the application fields of the resource allocation game model.
- **A co-evolution based algorithm** is proposed for obtaining the Nash equilibrium strategies based on practical action sets improved the feasibility of strategies analysis.
- Sufficient simulations based on **four real-world networks** proved the effectiveness of our proposed game.

Thank You