Journal of System Simulation

Volume 33 | Issue 2 Article 4

2-20-2021

Modeling of Adversarial Behavior on Road Network Based on Non-cooperative Game

Xiangyu Wei
1. China Huayin Ordnance Test Center, Huayin 714200, China; ;

Zhang Qi
2. College of Systems Engineering, National University of Defense Technology, Changsha 410073, China;

Follow this and additional works at: https://dc-china-simulation.researchcommons.org/journal

 Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Modeling of Adversarial Behavior on Road Network Based on Non-cooperative Game

Abstract

Abstract: The modeling of adversarial behavior is the key to the study of various military-like confrontation problems. Existing research mainly focus on the target domain, but in reality, many confrontation problems occur on the road network. *Combined with the network flow representation of adversarial behavior, a network adversarial game modeling framework based on non-cooperative games is proposed, and a novel road network confrontation problem—network evasion interdiction game is given based on the framework. Simulation experiments show that the new double oracle algorithm performs better than the original linear solution algorithm. Data experiments based on real road networks further verify the feasibility and scalability of the algorithm.* It shows that the road network interdiction game modeling framework based on non-cooperative game can better model adversarial behaviors on road network, and shows stronger robustness and better intelligence.

Keywords

non-cooperative game, network interdiction, adversarial behavior modeling, network evasion interdiction game

Recommended Citation

Wei Xiangyu, Zhang Qi. Modeling of Adversarial Behavior on Road Network Based on Non-cooperative Game[J]. Journal of System Simulation, 2021, 33(2): 271-279.

This paper is available in Journal of System Simulation: https://dc-china-simulation.researchcommons.org/journal/vol33/iss2/4

第 33 卷第 2 期 系统仿真学报© Vol. 33 No. 2
2021 年 2 月 Journal of System Simulation Feb. 2021

基于非合作博弈的路网对抗行为建模研究

魏翔宇[1]，张琪[2*]
(1. 中国华阴兵器试验中心，陕西 华阴 714200；2. 国防科技大学 系统工程学院，湖南 长沙 410073)

摘要：对抗行为建模是很多类军事竞争问题研究的关键。现有研究主要面向目标域，而现实中很多对抗问题发生在路网之上。对此，*结合对抗行为的网络流表示，提出了一种基于非合作博弈的路网对抗博弈建模框架，并给出了一个问题实例—网络逃避阻断博弈。仿真试验表明新的双启发式求解算法表现优于原始线性求解算法；基于现实路网的数据实验进一步验证了算法的可行性和可扩展性*，说明基于非合作博弈的路网对抗博弈建模框架可以更好地建模路网对抗行为，且表现出更强的鲁棒性和更好的智能性。

关键词：非合作博弈；网络阻断；对抗行为建模；网络逃避阻断博弈
中图分类号：TP391.9 文献标志码：A 文章编号：1004-731X (2021) 02-0271-09
DOI: 10.16182/j.issn1004731x.joss.20-0943

Modeling of Adversarial Behavior on Road Network Based on Non-cooperative Game

Wei Xiangyu[1], *Zhang Qi*[2*]
(1. China Huayin Ordnance Test Center, Huayin 714200, China;
2. College of Systems Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract: The modeling of adversarial behavior is the key to the study of various military-like confrontation problems. Existing research mainly focus on the target domain, but in reality, many confrontation problems occur on the road network. *Combined with the network flow representation of adversarial behavior, a network adversarial game modeling framework based on non-cooperative games is proposed, and a novel road network confrontation problem—network evasion interdiction game is given based on the framework. Simulation experiments show that the new double oracle algorithm performs better than the original linear solution algorithm. Data experiments based on real road networks further verify the feasibility and scalability of the algorithm.* It shows that the road network interdiction game modeling framework based on non-cooperative game can better model adversarial behaviors on road network, and shows stronger robustness and better intelligence.

Keywords: non-cooperative game; network interdiction; adversarial behavior modeling; network evasion interdiction game

引言

对抗行为建模[1-2]是很多类军事竞争问题研究的关键。智能的有适应能力的对手会实施监视行为观察和评估防守者的防御策略，然后采取攻击行为；而防守者的防御资源通常是有限的，无法对所关注的所有目标进行完全防御。无论是干扰敌方还是防护己方[3-4]。相对于一般的资源分配调度问题，受对手的决策行为空间影响，对抗行为建模问题更为复杂，其决策空间是自身行为空间与对手行为空

收稿日期：2020-11-30 修回日期：2020-12-24
第一作者：魏翔宇(1989-)，男，博士，工程师，研究方向为复杂系统建模与仿真，智能行为建模。E-mail：weixiangyu08@nudt.edu.cn
通讯作者：张琪(1988-)，男，博士，讲师，研究方向为复杂系统建模与仿真、智能行为建模。E-mail：zhangqiy123@nudt.edu.cn

Journal of System Simulation, Vol. 33 [2021], Iss. 2, Art. 4

间的笛卡尔乘积。

现有对抗行为模型的研究主要面向目标域，对抗的焦点在于攻击或防御特定的目标个体[5-6]。而现实中很多对抗问题发生在路网(Road Network)之上，例如军事活动中的后勤补给和兵力投送等行动都依托于路网进行。对于路网对抗问题，网络本身的结构和参数会影响对抗双方的行为，双方的对抗行为可以通过网络流模型进行建模。例如攻击者从初始节点移动到目的节点必须满足的路径可行性约束。具体的行为能力也可以用网络图的参数来量化，例如通过某段道路的时间可以用网络中边的长度来表示。

相对于目标域的对抗问题，路网对抗问题更为复杂。以孟买路网阻断问题[7]为例，攻击者的行为空间数量级超过 10^{18}。防守者的行为空间随可用防御资源的数目增长呈指数增长；攻击者的行为空间随网络的规模增长而指数增长。通常路网对抗问题是 NP-hard 的，对于如此规模的对抗问题，即便是采用最简单的线性求解方法也将面临巨大的计算挑战，因此必须考虑更紧凑的模型表示和可扩展良好的算法来建模基于路网的对抗行为。

图论中的网络流模型是一类优秀处理此类问题的研究工具。对抗双方的行为可以用最短路、最大流和最小费用流等网络流模型进行建模，例如通常攻击者的目的在于最优化网络的最短路或者最大流。博弈论是建模 Agent 之间竞争和合作行为的良好范式[8]，计算博弈理论为多 Agent 对抗行为建模提供了有效的数学框架，可以很好地建模攻防双方的对抗交互。尤其是 Stackelberg 领导者—追随者博弈模型更是被广泛地应用于建模面向目标域的攻防对抗行为。因此，采用网络流模型建模双方的行为，并采用非合作博弈处理行为之间的对抗交互，是路网对抗行为建模和求解的一个可行思路，据此，本体提出了一种采用非合作博弈的路网对抗行为建模框架。

Washburn 和 Wood[9]采用最短路模型结合两人零和博弈建模构建了一类网络阻断(network interdiction)问题模型，其中防守者试图通过选择最优的边策略来侦查攻击者，其阻断模型是二元模型，而在实际中，阻断行为受设备和人力等不确定因素的影响可能会失效。Pan 等[10]考虑了阻断行为是否成功的情形，拦截者试图通过监视网络中的节点和边来最大程度减少逃避者的最大逃避概率。肖开明等[11]研究了双目标最短路阻断问题，给出了一种子图求解算法并进一步分析了资源的饱和特性。Tsai 等[2]提出了一种线性求解模型 RANGER 来有效地计算机场安检中检查点设置的最佳边际分布，采用强 Stackelberg 均衡给出了概率化的日常调度混合策略。

1 基于非合作博弈的网络对抗行为建模框架

目前对于 Agent 对抗行为的研究多数是面向目标结构域的，本文重点处理基于路网的对抗行为建模问题。对于基于路网的对抗行为建模问题，攻防双方的对抗行为发生在道路网络 $G=(N,E)$上，其中 N 是节点集合，E 是网络的边的集合，对抗双方的行为能力可以用网络中边或节点的参数表示，例如道路的长度或者道路的旅行时间。

结合网络流模型，通常攻击者的目的在于最优化网络的最短路或者最大流等特定的目标函数，而防守者通过有限的防御资源对网络中的边或者节点进行阻断以最大化地干扰攻击者的目标。显然，对抗模型天然包含 max-min 冲突，对于此类场景下 Agent 的对抗行为所固有的强竞争性，通常采用非合作博弈理论进行数学建模。强对抗带来的直接结果就是博弈是零和的，即一方所得即一方所失。因此对于一般的路网对抗行为建模问题，可以采用博弈论的极大极小公式 $\min_{\boldsymbol{d}\in D}\max_{\boldsymbol{a}\in A}U_f(\boldsymbol{d},\boldsymbol{a})$ 来表示双方的最优策略，其中 $\boldsymbol{d}$ 和 $\boldsymbol{a}$ 分别表示防守者和攻击者的策略。这里的 $\boldsymbol{d}$ 既可以是纯策略也可以混合策略。如果 $\boldsymbol{d}$

https://dc-china-simulation.researchcommons.org/journal/vol33/iss2/4
DOI: 10.16182/j.issn1004731x.joss.20-0943

Wei and Qi: Modeling of Adversarial Behavior on Road Network Based on Non-coo

是纯策略，可以表示防守者的资源分配策略，那么此时 D 是所有可行防御资源分配策略的集合；如果 $\boldsymbol{d}$ 是混合策略，表示防守者在所有可能资源分配策略上的概率分布选择，此时 D 就对应线性规划的单纯型。同样对于攻击者来说，攻击策略 $\boldsymbol{a}$ 可以网络流来进一步抽象表示，此时对应的 A 就是所有满足流约束的可行网络流的集合。

结合网络流和非合作博弈，本文提出一个基本的网络结构域下的对抗行为建模框架，并依据此框架建模一个新的路网对抗问题，并进一步给出新的对抗博弈模型的求解方法。

如图 1 所示，对于给定的具体的网络对抗行为建模问题实例，首先需要确定双方对抗行为的博弈策略表示。结合网络拓扑和网络流约束，可以给出对抗双方行为的纯策略表示，然后根据博弈中混合策略的定义，可以进一步给出对抗双方的效用表示。由对抗问题的零和假设，这里实际上只需要给出一方的策略表示。随后就是选择合适的解概念并确定博弈的均衡解是否存在。本文选用强 Stackelberg 均衡作为对抗博弈的解概念，其能更好地反映博弈的对抗过程。对于零和博弈，强 Stackelberg 均衡是必定存在的，因为先行的一方(在 Stackelberg 博弈中称为领导者)至少可以承诺实施纳什均衡策略。如此根据强 Stackelberg 均衡的定义就可以构建对抗行为博弈模型的均衡求解公式，如图 1 左虚线框所示，如此就完成了对抗行为的博弈模型构建。

构建博弈模型后，问题的重点在于如何求解。基于零和博弈中强 Stackelberg 均衡与极大极小均衡的等价性，可以采用极大极小定理公式来求解。如前文所述，由于博弈模型的复杂度是 NP-hard 的，难以直接在原始问题上利用线性公式求解。一个可行的思路是采用迭代式求解算法，以小规模博弈的不断迭代增长逼近原始大规模问题的一次性求解。此外，针对特定问题的网络流表示，例如最短路或者最大流公式等，可以进一步应用一些启发式算法对算法进行进一步的加速，提高每次迭代的求解效率。如此，如图 1 右虚线所示，就完成了非合作博弈模型的求解。

下面以一个新的路网对抗为问题实例来展示如何基于非合作博弈建模基于路网的对抗行为。

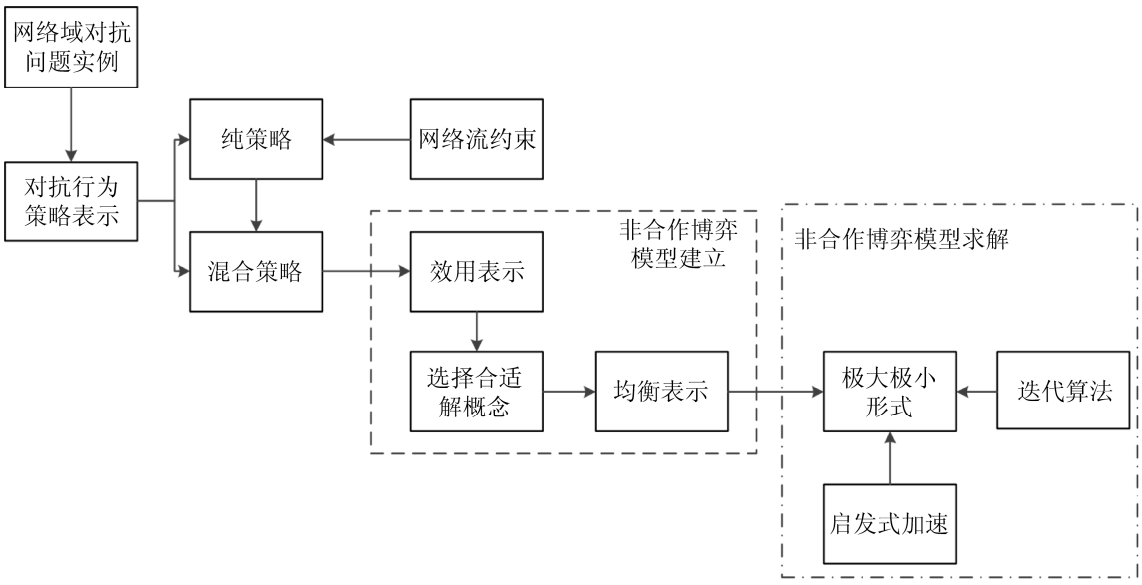


图 1 基于非合作博弈的网络对抗行为建模框架

Fig. 1 Adversarial behavior modeling framework on network structured domains based on non-cooperative games

Published by Journal of System Simulation, 2021

Journal of System Simulation, Vol. 33 [2021], Iss. 2, Art. 4

2 网络逃避阻断博弈问题

基于 Pan 和 Morton 的最小化最大可靠路径网络阻断模型[10]，考虑路网对抗行为中存在的不确定性和不同目标的优先级，构建如下问题：攻击者企图通过路网攻击网络中的某个目标节点，不同的目标节点的收益值(Payoff)不同。攻击者的攻击路线受网络可行路径约束，其目的在于最大化攻击的期望收益，即成功到达目标节点的概率与相应目标节点的收益的乘积。由零和假设，显然防守者的目的在于最小化攻击者的最大期望收益。为了行文方便，记这个新的道路网络上的逃避阻断博弈问题为 NEIG (Network Evasion Interdiction Game)。

路网逃避阻断博弈问题 NEIG 定义在网络图 $G(N,E)$上，其中节点集合为 N，有向边集合为 E，具体的符号和标记如表 1 所示。

表 1 网络逃避阻断问题符号标记系统
Tab. 1 Notations of NEIG

主要参数	含义
$G(N,E)$	NEIG 网络图
R	防御资源总预算
p_{uv}	攻击者在边(u,v)上的原始逃脱概率
q_{uv}	阻断边(u,v)后攻击者的逃脱概率
c_{uv}	阻断边(u,v)所需要消耗的资源
$\boldsymbol{x}^i \in \{0,1\}^{\lvert E\rvert}$	防守者的纯策略
$X=\{\boldsymbol{x}^i\}$	防守者的可行纯策略集合，
$\boldsymbol{d}$	防守者的混合策略
$\boldsymbol{y}^j \in \{0,1\}^{\lvert E\rvert}$	攻击者的纯策略
$Y=\{\boldsymbol{y}^j\}$	攻击者的可行路径集合
$\boldsymbol{a}$	攻击者的混合策略

网络逃避阻断博弈是一类典型的 Stackelberg 博弈，防守者对应博弈的领导者先采取行动，在网络的边上布置防御资源；攻击者对应博弈的追随者随后行动，其选择的攻击路线受网络可行路径约束，其目的在于最大化攻击的期望收益。

基于前文给出的建模框架，进一步确定对抗双方行为的策略表示。

2.1 对抗行为的策略表示

参照表 1，防守者的纯策略是防御资源在网络边上的分配，定义为向量 $\boldsymbol{x}^i \in \{0,1\}^{|E|}$，表示防守的第 i 个纯策略。如果边(u,v)受到阻断，即防守者分配相应的防御资源，那么则 $x_{uv}^i=1$，否则 $x_{uv}^i=0$。防守者的可行纯策略集合记为 $X=\{\boldsymbol{x}^i \mid i=1,...,n\}$。防守者的混合策略记为 $\boldsymbol{d}=\langle d_X \rangle$，也即 X 上的概率分布，其中 d_i 表示执行第 i 个纯策略 $\boldsymbol{x}^i$ 的概率。防守者需要消耗 c_{uv} 单位的资源来阻断边(u,v)，其防御资源总预算为 R。

相对于防守者仅受有限资源预算的约束，攻击者的纯策略 $\boldsymbol{y}^j \in \{0,1\}^{|E|}$ 必须满足图的可行路径约束，简称为路径 j。对于任意的边(u,v)，如果路径 j 经过该边，则 $y_{uv}^j=1$；否则 $y_{uv}^j=0$。攻击者的所有可行路径集合用 Y 表示，对于集合 $Y=\{\boldsymbol{y}^j \mid j=1,2,...,m\}$ 中的路径，需要满足网络流可行路径约束：

$$Y=\{\boldsymbol{y}\geqslant 0:\sum_{(u,v)\in E} y_{uv}-\sum_{(u,v)\in E} y_{vu}=\boldsymbol{B},\forall u\in N\} \quad (1)$$

其中 $\boldsymbol{B}=[1,0,...,-1]_{|N|\times 1}$。

2.2 效用表示

基于攻击者策略的路径表示，可以得出攻击者的效用表示。记路径 j 上对应的目标节点为 t_j，其收益为 ω_{t_j}，那么攻击者的效用就是攻击者沿路径 j 成功逃脱的概率与目标节点 t_j 收益 ω_{t_j} 的乘积。这里假设攻击者在各条边上的逃脱概率相互独立，那么攻击者通过路径 j 成功逃脱的概率就是路径 j 上各个边对应的概率乘积。

为了方便说明前文的定义和符号标记，以一个网络逃避阻断博弈实例为说明。如图 2 所示，网络中有 2 个源节点 s_1 和 s_2，2 个目标节点 t_1 和 t_2，图中各边旁边的标记对应 2 种逃脱概率$(p_{uv},\ q_{uv})$，以红色和蓝色分别标记 2 条攻击者的可行攻击路径，红色路径为 $s_1\rightarrow s_2\rightarrow t_1$，蓝色路径为 $s_2\rightarrow 2\rightarrow 3\rightarrow t_2$。

https://dc-china-simulation.researchcommons.org/journal/vol33/iss2/4
DOI: 10.16182/j.issn1004731x.joss.20-0943

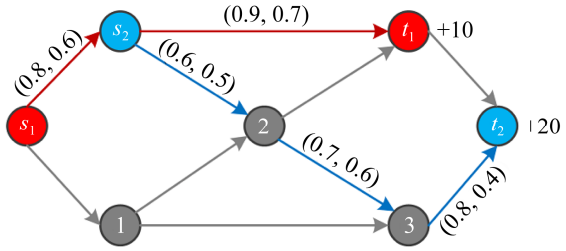


图 2 一个简单的网络逃避阻断博弈实例
Fig 2. Simple example of network evasion interdiction game

如下，考虑防守者的 3 种阻断方案；

(1) 未施加阻断行为；

(2) 只阻断蓝色路径(边 $(3,t_2)$)；

(3) 同时阻断红色和蓝色路径(边$(3,\ t_2)$和(s_2, t_1))。

3 种不同阻断策略下，攻击者对应的期望收益如表 2 所示。

表 2 红蓝路径在不同防守者阻断策略下的期望收益
Tab. 2 Expected payoffs of red and blue paths under different interdiction strategies of defender

路径	ω_t	期望收益		
		0	$\{(3, t_2)\}$	$\{(3, t_2), (s_2, t_1)\}$
红色	20	7.20	7.20	5.60
蓝色	10	6.72	3.84	3.84

攻击者的目标在于最大其效用，即最大化攻击者沿路径成功逃脱的概率与路径对应目标节点的收益的乘积。根据网络逃避阻断博弈的零和假设，相应的防守者的目标就是最小化攻击者的效用。

攻击者的效用可以用公式(2)来表示:

$$U_a(\boldsymbol{x}^i,\boldsymbol{y}^j)=\omega(t_j)\left\{\prod_{(u,v)\in y^j}\left[p_{uv}(1-x_{uv}^i)+q_{uv}x_{uv}^i\right]\right\}=$$
$$\omega(t_j)\left\{\prod_{(u,v)\in E}\left[p_{uv}(1-x_{uv}^i)+q_{uv}x_{uv}^i\right]^{y_{uv}^j}\right\}=$$
$$\omega(t_j)\Phi(\boldsymbol{x}^i,\boldsymbol{y}^j) \tag{2}$$

式中：$\Phi(\boldsymbol{x}^i,\boldsymbol{y}^j)=\prod_{(u,v)\in E}[p_{uv}(1-x_{uv}^i)+q_{uv}x_{uv}^i]^{y_{uv}^j}$，表示在给定防守者的资源分配策略 $\boldsymbol{x}^i$ 情形下攻击者沿路径 j(用 $\boldsymbol{y}^i$ 表示)成功逃脱的概率；$\omega(t_j)$为路径 j 上对应的目标节点为 t_j 的收益。由各边逃脱概率的独立性假设，攻击者通过路径 j 成功逃脱的概率就是路径 j 上各个边对应的概率乘积。在公式(2)中，$\forall(u,v)\in \boldsymbol{y}^j$，$p_{uv}(1-x_{uv}^i)+q_{uv}x_{uv}^i$ 可以用来表示给定防守者纯策略 $\boldsymbol{x}^i$ 情况下，攻击者通过边(u,v)成功逃脱的概率。具体而言，如果边(u,v)被防守者阻断，即 x_{uv}=1，则有 $p_{uv}(1-x_{uv})$=0，此时攻击者通过边(u,v)成功逃脱的概率为 q_{uv}；否则攻击者通过边(u,v)成功逃脱的概率为 p_{uv}。由零和假设，相应的，防守者的效用$U_d(\boldsymbol{x}^i,\boldsymbol{y}^j)=-U_a(\boldsymbol{x}^i,\boldsymbol{y}^j)$。

用向量 $\boldsymbol{d}$ 表示防守者的混合策略，那么给定攻击者的纯策略 y^j，防守者采用混合策略 $\boldsymbol{d}$ 的效用可以由公式(3)计算得出，即：

$$U_d(\boldsymbol{d},\boldsymbol{y}^j)=\sum_{i=1}^{n}U_d(\boldsymbol{x}^i,\boldsymbol{y}^j)d_i \tag{3}$$

2.3 均衡表示

网络逃避阻断博弈属于 Stackelberg 博弈，符合领导者-追随者博弈范畴。此类问题中应用最广泛的解概念是强 Stackelberg 均衡。

令 $y(\boldsymbol{x})$表示攻击者针对防守者混合策略 $\boldsymbol{x}$ 的最优应对纯策略，那么策略组合 $\langle \boldsymbol{x}^*,\boldsymbol{y}^*\rangle$ 构成强 Stackelberg 均衡，当其满足以下 2 个条件：

(1) $\boldsymbol{y}^*=y(\boldsymbol{x}^*)$；

(2) 对于所有的 $\boldsymbol{x}\in X$ 都有 $U_d(\boldsymbol{x}^*,\boldsymbol{y}^*)\geqslant U_d(\boldsymbol{x},y(\boldsymbol{x}^*))$。

由于在零和假设中，强 Stackelberg 均衡、纳什均衡和极大极小均衡是等价的，因此强 Stackelberg 均衡可由公式(4)得出：

$$\min_{\boldsymbol{x}\in X}\max_{\boldsymbol{y}\in Y}U_a(\boldsymbol{x},\boldsymbol{y}) \tag{4}$$

式中：$U_a(\boldsymbol{x},\boldsymbol{y})=-U_d(\boldsymbol{x},\boldsymbol{y})$。

由 von Neumann 的最小最大定理，可以得出：

$$\min_{\boldsymbol{x}\in X}\max_{\boldsymbol{y}\in Y}U_a(\boldsymbol{x},\boldsymbol{y})=\max_{\boldsymbol{y}\in Y}\min_{\boldsymbol{x}\in X}U_a(\boldsymbol{x},\boldsymbol{y}) \tag{5}$$

极大极小定理的直接推论就是强 Stackelberg 均衡和纳什均衡的等价性，即 $\boldsymbol{x}^*$是 $\boldsymbol{y}^*$的最优应对的同时，$\boldsymbol{y}^*$也是针对 $\boldsymbol{x}^*$的最优应对。这就启发我们可以用双启发式算法(a double oracle algorithm)来求解博弈的均衡解，即给定一个初始解，通过不断生成双方的最优应对，直至算法达到收敛。

Published by Journal of System Simulation, 2021

Journal of System Simulation, Vol. 33 [2021], Iss. 2, Art. 4

3 双启发式求解算法

基于对抗双方最优应对的相互迭代，双启发式求解算法，可以通过迭代逐步增大博弈的规模，无需直接求解原始的完全大规模问题。

3.1 极大极小公式

根据极大极小定理和强 Stackelberg 均衡的定义，给定防守者的纯策略集合 X 和攻击者的纯策略集合 Y，可以通过求解公式(6)获得对抗双方的均衡混合策略。

$$\max_{U^*,\boldsymbol{d}} \quad U^* \\ \text{s.t.}\begin{cases} U^* \leqslant -U(\boldsymbol{d},\boldsymbol{y}^j) \quad \forall j=1,2,...,m \\ \boldsymbol{I}^{\mathrm{T}}\boldsymbol{d}=1 \\ \boldsymbol{d}\in[0,1]^{|E|} \end{cases} \tag{6}$$

该线性规划可以用来直接求解对抗行为的原始博弈模型，但是，如前文所述，模型的变量和约束随问题的规模 $(|X|,|Y|)$ 指数增长，线性规划本身会面临可扩展性和鲁棒性的问题，在实际求解中难以应用到现实路网规模的问题。公式(6)是求解网络逃避阻断博弈模型的核心，记为函数 $Core(X,Y)$。

3.2 算法代码

基于 $Core(X,Y)$，双启发式求解算法伪代码如算法 1 所示。

算法 1 网络逃避阻断博弈的双启发式求解算法

输入： 问题实例

1：初始化 X'：随机将 R 资源分配到 E 上
2：初始化 Y'：随机生成一条可行 s–t 路径
3：repeat
4：　　$(\boldsymbol{d},\boldsymbol{a}) \leftarrow Core(X',Y')$
5：　　$\boldsymbol{x}^* \leftarrow DBR(\boldsymbol{a})$
6：　　$X' \leftarrow X' \cup \{\boldsymbol{x}^*\}$
7：　　$\boldsymbol{y}^* \leftarrow ABR(\boldsymbol{d})$
8：　　$Y' \leftarrow Y' \cup \{\boldsymbol{y}^*\}$
9：until convergence
10：return $\langle \boldsymbol{d},\boldsymbol{a} \rangle$

双启发式算法的核心思路在于迭代式地维护一个原始博弈模型的受限博弈(restricted game)，通过每一轮对抗双方针对彼此的最优应对逐步地扩展受限博弈的规模，直至其与原始博弈存在相同的均衡。显然算法是必定收敛的，因为最坏的情况下受限博弈至多与原始模型相同。实际上，受限博弈的规模往往远小于原始博弈的规模，这就使得算法可以更快速地求得均衡解，从而提高了算法的求解效率。

记受限博弈中防守者和攻击者的行为决策空间分别为 X'和 Y'，首先随机生成攻击者和防守者的任意可行纯策略来初始化 X'和 Y'，然后求解 $Core(X',Y')$及其对偶就获得受限博弈的均衡解，原始解 $\boldsymbol{d}$ 和对偶解 $\boldsymbol{a}$ 必定满足受限博弈的强 Stackelberg 均衡，然而这一组解却不一定是原始问题的强 Stackelberg 均衡解。此时转到原始博弈问题上来，分对抗双方别针对 $Core(X',Y')$所产生的原始解 $\boldsymbol{d}$ 和对偶解 $\boldsymbol{a}$ 产生各自的最优应对纯策略，再将产生的纯策略纳入受限博弈(X', Y')中来，以进一步扩充受限博弈逼近原始博弈，并引导受限博弈的均衡解逐步收敛到原始博弈的强 Stackelberg 均衡。这其中的关键就是如何产生攻击者和防守者各自的最优应对。记攻击者针对防守者混合策略 $\boldsymbol{d}$ 产生最优应对的启发式子问题为 ABR($\boldsymbol{d}$) (Attacker Best Response)，ABR 子问题可以生成一条攻击者的可行路径；记防守者针对攻击者混合策略 $\boldsymbol{a}$ 产生最优应对的启发式子问题为 DBR($\boldsymbol{a}$) (Defender Best Response)，DBR 子问题可以生成一种防守者的资源分配方案。在迭代过程中，一旦子问题 ABR 和 DBR 所产生的最优应对已经存在于受限博弈(X',Y')中，算法终止返回相应的结果，当前的策略组合 $\langle \boldsymbol{d},\boldsymbol{a} \rangle$ 就是网络逃避阻断博弈的强 Stackelberg 均衡解。对于混合策略，其行为空间可以看作决策空间上的单纯型，算法的收敛性可以用 von Neumann 极大极小定理证明。

3.3 防守者的启发式，DBR

DBR 是防守者针对攻击者混合策略 $\boldsymbol{a}$ 产生最

https://dc-china-simulation.researchcommons.org/journal/vol33/iss2/4
DOI: 10.16182/j.issn1004731x.joss.20-0943

优应对的启发式子问题，记为 DBR(***a***)，用以生成新的防守者资源分配方案纯策略。

根据图的广义网络流[12]表示，DBR(***a***)可以用如下的双层规划来构建，如公式(7)所示。

$$\min_{\boldsymbol{x}} \max_{f,f'} \sum_{j=1}^{m} a_j \left(\sum_{(t_j,t)} \omega_{t_j} f_{t_j t}^{j} \right)$$

$$\text{s.t.} \begin{cases} \sum_{(u,v)\in E} (f_{uv}^{j} + f_{uv}^{'j}) - \sum_{(u,v)\in E} (p_{vu} f_{vu}^{j} + p'_{vu} f_{vu}^{'j}) = B_u & \forall u \in N \setminus t, \forall j \\ f_{uv}^{j} \leqslant 1 - x_{uv} & \forall (u,v) \in E, \forall j \\ f_{uv}^{j} + f_{uv}^{'j} \leqslant y_{uv}^{j} & \forall (u,v) \in E, \forall j \\ \sum_{(u,v)\in E} c_{uv} x_{uv} \leqslant R & \\ x_{uv} \in \{0,1\} & \\ f_{uv}^{j}, f_{uv}^{'j} \geqslant 0 & \forall j \end{cases} \tag{7}$$

目标函数为最小化攻击者的最大期望效用。攻击者的期望效用是攻击者在策略空间 Y 各路径上的效用与混合策略 ***a*** 的乘积。参考广义网络流，约束条件 1 对应广义的节点流平衡约束，流入边的网络流需要按照攻击者的成功逃脱概率 p_{vu}(或 p'_{vu})加以调整。B 表示除了目标节点之外其他节点的网络流供给或需求，对于源节点有 B_s=1，其他节点 B_u=0。约束条件 2 保证所有被防守者所阻断的边上的网络流满足 f_{uv}^{j}=0，也被称为截流约束。如果边 (u,v)受到阻断(x_{uv}=1)，那么对于 $\forall j$，都有 f_{uv}^{j}=0，这就迫使网络流只能选择以 $f_{uv}^{'j}$ 的方式通过被阻断边，显然 $f_{uv}^{'j}$ 对应了更低的逃脱概率。约束条件 3 保证了非 0 的网络流仅存在于攻击者所选择的路径上。约束条件 4 表示总的防御资源预算约束。约束条件 5 和 6 给定了变量的取值范围。

3.4 攻击者启发式，ABR

ABR 是攻击者针对防守者混合策略 ***d*** 产生最优应对的启发式子问题，记为 ABR(***d***)，用以生成新的攻击者逃避路径选择纯策略。

ABR(***d***)可以用式(8)所示的数学公式构建。

目标函数为考虑防守者的混合策略 ***d*** 的情况下，最大化攻击者选择路径 ***y*** 的期望效用。约束条件 1 为可行路径的节点平衡约束，即保证 ***y*** 是一条简单无环路径。同样的，参考广义网络流，约束条件 2 对应广义的网络流节点平衡约束，即网络的流受到防守者纯策略 $\boldsymbol{x}^i$ 的调节。约束条件 3 保证网络流被限制在攻击者的逃避路径 ***y*** 上。约束条件 4 中，若边为防守者所阻断(x_{uv}^{i}=1)，则有 f_{uv}^{i}=0，因此也被成为非截流条件。约束条件 5 和 6 给定了变量的取值范围。

$$\max_{\boldsymbol{y},f,f'} \sum_{i=1}^{n} d_i \left(\sum_{(u,t)\in E} \omega_u f_{ut}^{i} \right)$$

$$\text{s.t.} \begin{cases} \sum_{(u,v)\in E} y_{uv} - \sum_{(u,v)\in E} y_{uv} = b_u & \forall u \in N \setminus t \\ \sum_{(u,v)\in E} (f_{uv}^{i} + f_{uv}^{'i}) - \sum_{(u,v)\in E} (p_{vu} f_{vu}^{j} + p'_{vu} f_{vu}^{'j}) = \mathrm{b}_u & \forall u \in N \setminus t, \forall j \\ f_{uv}^{i} + f_{uv}^{'i} \leqslant y_{uv} & \forall (u,v) \in E, \forall j \\ f_{uv}^{i} \leqslant 1 - x_{uv}^{i} & \forall (u,v) \in E, \forall j \\ y_{uv} \in \{0,1\}^{|E|} & \\ f, f' \geqslant 0 & \end{cases} \tag{8}$$

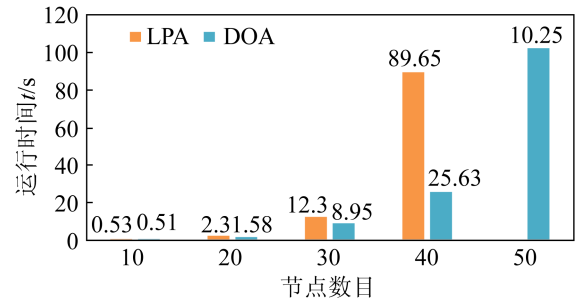
4 仿真实验

仿真实验采用随机 Waxman 图模型生成模拟的路网拓扑。$|N|$ 个节点的 Waxman 图模型用 $SimRoadNetwork(\lambda,\alpha,\beta,domain)$ 来表示，参数的具体含义可以参考文献[13]。随机生成初始逃脱概率 $p_{uv} \in [0.5,1)$；阻断后逃脱概率随机下降 $(0,0.5]$。防御资源总预算 $R = |\mu| E||$ 向上取整，通过调节参数 μ 设置预算总量。

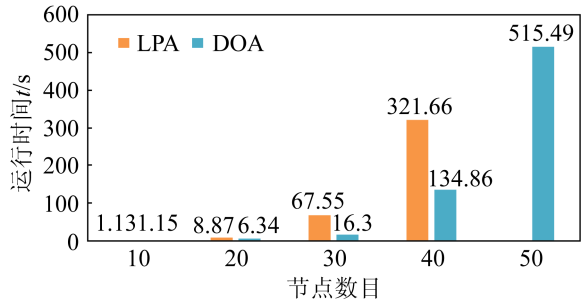
通过合理调整参数 α 和 β 可以获得期望的平均度 D，该参数可以控制网络中连边的数量。实验分别选取 3 个平均度数，即 $D \in \{2.7, 2.86, 3.0\}$，其中根据 Gastner 等的研究，D=2.86 最符合实际路网拓扑[14]。

以求解原博弈的直接线性规划求解方法为基准算法，记为 LPA 算法[15]，记双启发式算法为 DOA，对比在不同平均度下随节点数目增大情况下的算法运行时间，实验结果如图 3 所示，在柱状图中 x 轴为网络节点的数目，y 轴为算法的求解时间，

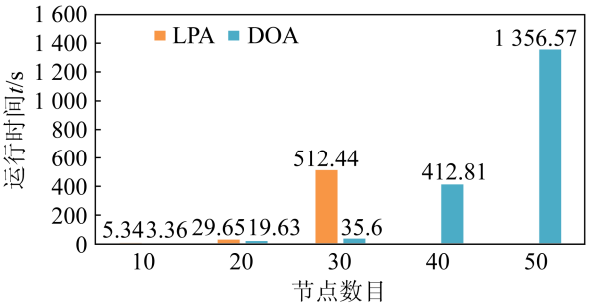
橙色为 LPA 算法的平均求解时间，蓝色为 DOA 算法的平均求解时间。注意如果算法无法在 3 600 s 内给出结果，则认为算法无法求解相应场景。



(a) D=2.7



(b) D=2.86



(c) D=3.0

图 3 LPA 算法和 DOA 算法的可扩展性对比

Fig. 3 Scalability comparison between LPA algorithm and DOA algorithm

LPA 算法和 DOA 算法的求解能力均受网络规模的影响，随着网络规模的增长，2 种算法的求解时间均有显著的上升，但是 DOA 算法的上升趋势要显著慢于 LPA 算法，说明基于迭代的双启发式算法 DOA 表现明显优于基于线性规划的 LPA 算法。LPA 算法的可扩展性较差，随着节点数目的增加，算法的求解时间急剧上升，甚至难以在有效时间内给出均衡解，以 D=3.0 实验为例，节点数目超过 30 后 LPA 算法无法在规定时间内求解。综上所述，论文提出的 DOA 算法求解性能要优于原始算法，并且表现出更好的可扩展性。

参数 D 直接影响网络中连边的密度，对比图 3 中 3 个子图，以 40 节点为例，随着 D 取值的增大，网络中的连边数量增多，LPA 算法和 DOA 算法所发的求解时间均会增长，这说明，网络中边的数量显著影响算法的求解时间。如图 3(c)所示，LPA 算法无法求解如此规模的网络，进一步验证了 DOA 算法的可扩展性。

考虑现实中的道路网络，如图 4 所示，芝加哥路网共包含 933 节点和 2 950 条边。网络假设为无向图，源节点和目标节点如图 4 所示。网络中各边的长度取道路实际长度的整数近似。阻断的资源消耗与点的度数成正比。给定不同的防御资源预算，对比 LPA 算法和 DOA 算法的表现，如表 3 所示。显然 LPA 算法在现实规模问题的处理上表现一般，而基于非合作博弈的路网对抗行为建模框架及求解算法能够更好地处理现实规模的网络问题。

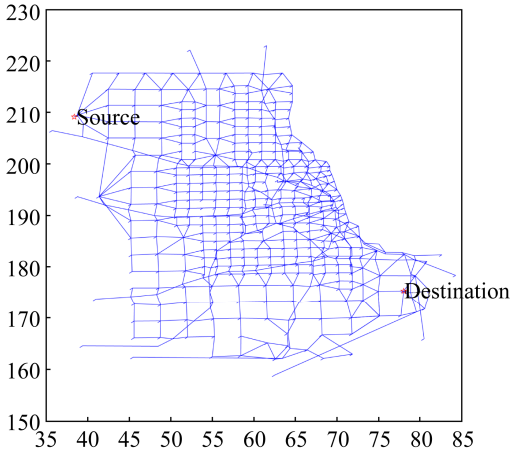


图 4 芝加哥路网的 matlab 模拟

Fig. 4 Matlab simulation of Chicago sketch road network

表 3 芝加哥路网的计算结果

Tab. 3 Calculation results on Chicago road network

No.	资源预算	求解时间/s	
		LPA	DOA
1	5	5.46	4.58
2	10	36.52	20.54
3	100	—	455.63

https://dc-china-simulation.researchcommons.org/journal/vol33/iss2/4
DOI: 10.16182/j.issn1004731x.joss.20-0943

5 结论

攻防对抗是军事安全领域长期面临的一个关键问题，尤其是在可以用图或者网络进行建模的领域，如何处理智能化适应性的对手行为以及网络流特性对 Agent 之间交互的影响，更智能地建模攻防双方的对抗行为是亟需解决的一个问题。本文基于非合作博弈，从网络流的角度提出了一种新的网络对抗行为建模的框架。以一个新的路网对抗问题—网络逃避阻断博弈(NEIG)为实例，展示了如何基于非合作博弈建模基于路网的对抗行为。NEIG 模型充分考虑竞争双方的对抗行为所带来的不确定性，能够为防守者提供鲁棒性更强的混合防守策略。原始的模型求解算法难以处理大规模的网络问题，本文提出了一种基于双启发式的迭代式求解算法，并采用仿真试验验证了新的双启发式求解算法表现优于原始的线性规划求解算法。此外，基于现实路网的数据实验进一步验证了算法的可行性和可扩展性。本文提出的基于非合作博弈的路网对抗博弈建模框架建模路网对抗行为是可行的，并且表现出更强的鲁棒性和更好的智能性。

参考文献：

[1] Zhang J, Zhuang J, Behlendorf B. Stochastic Shortest Path Network Interdiction with a Case Study of Arizona–Mexico Border[J]. Reliability Engineering & System Safety (S0951-8320), 2018, 179: 62-73.

[2] Tsai J, Qian Y, Vorobeychik Y, et al. Tambe, Bayesian Security Games for Controlling Contagion[C]. 2013 International Conference on Social Computing. Alexandria, VA: IEEE Computer Society, 2013: 33-38.

[3] Washburn A. Network Interdiction[M]. Washburn A. Two-Person Zero-Sum Games. Dordrecht: Springer, 2014: 123-141.

[4] Sinha A, Fang F, An B, et al. Stackelberg Security Games: Looking Beyond a Decade of Success[C]. Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18). Stockholm, Sweden: IJCAI, 2018: 5494-5501.

[5] Pita J, Jain M, Marecki J, et al. Deployed ARMOR Protection: the Application of a Game Theoretic Model for Security at the Los Angeles International Airport[C]. 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006). Estoril, Portugal: DBLP, 2008.

[6] Paruchuri P, Tambe M, Fernando Ordóez, et al. Security in Multiagent Systems by Policy Randomization[C]. 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006). Hakodate, Japan: ACM, 2006.

[7] Jain M, Korzhyk D, Vanek O, et al. A Double Oracle Algorithm for Zero-sum Security Games on Graphs[C]. 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011). Taipei, Taiwan: International Foundation for Autonomous Agents and Multiagent Systems, 2011: 1-3.

[8] Fudenberg D, Tirole J. Game Theory[J]. Economica (S0013-0427), 1992, 60(238): 841-846.

[9] Washburn A, Wood K. Two-person Zero-sum Games for Network Interdiction[J]. Operations Research (S0030-364X), 1995, 43(2): 243-251.

[10] Pan F, Morton D P. Minimizing a Stochastic Maximum-reliability Path[J]. Networks (S0028-3045), 2010, 52(3): 111-119.

[11] Xiao K, Zhu C, Zhang W, et al. The Bi-objective Shortest Path Network Interdiction Problem: Subgraph Algorithm and Saturation Property[J]. IEEE Access (S2169-3536), 2020, 8: 1-1.

[12] Liu C. Generalized Network Flow Model with Application to Power Supply-demand Problems[R]. United States: Operations Research Center, 1982.

[13] Waxman B M. Routing of multipoint connections[J]. IEEE J. Select. Areas Commun (S0733-8716), 1988, 6(9): 1617-1622.

[14] Gastner M T, Newman M E J. The Spatial Structure of Networks[J]. The European Physical Journal B-Condensed Matter and Complex Systems (S1434-6028), 2006, 49(2): 247-252.

[15] Nguyen T H, Kar D, Brown M, et al. Towards a Science of Security Games[M]. Mathematical Sciences with Multidisciplinary Applications. Cham, Switzerland: Springer International Publishing, 2016.