面向化工园区有害气体泄漏的博弈巡逻策略

陈茵[1]，张立泽[2]，帅国华[3]，陈黎丽[2]，王震[2,4]
（1. 中国电信股份有限公司广西分公司，广西 南宁 530022；
2. 杭州电子科技大学网络空间安全学院，浙江 杭州 310018；
3. 中国（广西）自由贸易试验区钦州港片区应急管理局，广西 钦州 535008；
4. 杭州电子科技大学卓越学院，浙江 杭州 310018）

摘　要：在推进智慧城市一体化的背景下，智慧化工园区的建设步伐正在逐步加快。近些年，我国化工行业事故频发，给公共安全造成了巨大的破坏与损失。因此，如何提升化工园区安全管理和应急响应能力是目前亟待解决的问题。针对该类安全问题，提出面向化工园区有害气体泄漏的博弈巡逻策略。首先，利用对流扩散模型对有害气体泄漏的过程进行描述。其次，引入博弈论对巡逻问题中攻防双方的对抗过程进行建模，将防守者对安全事故的响应时间与其收益关联。然后，提出基于多线性规划的 GGC 算法，对该博弈模型进行求解。最后，在 3 种不同规模的真实化工园区案例场景下，将该博弈模型与其他两种基础方法进行收益对比。结果表明，该模型能够有效提升防守者收益，降低攻击者收益。

关键词：化工园区；气体泄漏；博弈巡逻；博弈论；响应时间
中图分类号：TP309.2
文献标志码：A
doi: 10.11959/j.issn.2096−6652.202338

Game patrol strategy for hazardous gas leakage in chemical parks

CHEN Yin[1], ZHANG Lize[2], SHUAI Guohua[3], CHEN Lili[2], WANG Zhen[2,4]
1. China Telecom Corporation Limited, Guangxi Branch, Nanning 530000, China
2. School of Cyberspace, Hangzhou Dianzi University, Hangzhou 311100, China
3. Qinzhou Port Emergency Bureau in China Guangxi Pilot Free Trade Zone, Qinzhou 535000, China
4. ZhuoYue Honors College, Hangzhou Dianzi University, Hangzhou 311100, China

Abstract: In the context of promoting the integration of smart cities, the pace of construction of smart chemical parks is gradually accelerating. Since accidents in China's chemical industry have occurred frequently in recent years, causing great damage and loss to public safety, improving the safety management and emergency response capability of chemical parks is an urgent need to be solved. For this kind of safety problems, this paper proposed a game patrol strategy for hazardous gas leakage in chemical parks. Firstly, the convective diffusion model was used to describe the process of hazardous gas leakage. Secondly, game theory was introduced to model the confrontation process between the attacking and defending parties in the patrol problem, and the response time of the defenders to a safety incident was correlated with its benefit. Then, a multilinear programming-based GGC algorithm was proposed to solve this game model. Finally, the gains of the game model in this paper were compared with the other two basic methods in the scenarios of three real chemical park case with different sizes. The results show that the model can effectively improve the gains of the defender and reduce the gains of the attacker.

Key words: chemical park, gas leakage, game patrol, game theory, response time

收稿日期：2023−03−14；**修回日期：**2023−05−23
通信作者：王震，wangzhen@hdu.edu.cn

0　引言

随着物联网、云计算、移动互联网、大数据和人工智能等新一代信息技术的飞速发展，智慧化工园区的建设步伐正在逐步加快，并积极融入智慧城市建设的大潮。智慧化工园区不仅能提升园区的运营管理能力，还能增强对重大安全风险的防控能力。然而，近些年，化工园区安全事故时有发生，部署巡逻队是预防安全问题最有效、最直接的手段[1]。园区管理者急需一个具有应急响应能力、成本效益高的巡逻策略。

目前，在实际化工园区的巡逻过程中，以固定巡逻策略和完全随机策略为主[2]。固定巡逻策略指的是巡逻队采用固定巡逻路线对园区内的道路进行多次巡逻。然而，理性攻击者可以通过观察巡逻队的活动来获取巡逻策略，并在特定时机对化工厂发起攻击，这使得巡逻失去原有意义。完全随机巡逻策略指的是巡逻队的巡逻路线是完全随机的，对每个化工厂巡逻的概率都是相等的。在这种情况下，理性攻击者往往会倾向于攻击具有最大价值的化工厂，即使冒着被发现的风险。由此可见，上述两种巡逻策略的巡逻效果十分有限，无法很好地对抗理性攻击者。针对该类攻击者，相关研究工作利用博弈论方法设计巡逻策略以实现安全巡逻的目标。

在化工园区巡逻领域中，文献[3]利用博弈论相关知识设计管道系统内的安全巡逻策略，实现了对高风险管道系统的覆盖。文献[2]提出基于博弈论的安全巡逻策略调度方法，利用Stackelberg均衡得到的巡逻策略表现出较好的性能。文献[4]提出基于博弈论的多辆巡逻车合作的巡逻策略，通过对化工园区进行分区，提高巡逻的整体收益。在化工园区巡逻过程中，巡逻队除了预防安全事故，还应在发生安全事故后及时响应。虽然博弈论能够有效构建攻防双方的收益模型并找到均衡解，但是巡逻策略还应该考虑到巡逻的应急响应情况，也就是巡逻的响应时间。其中，响应时间是指巡逻队发现安全事故后前往事发地的时间。当前，在化工园区巡逻工作中，研究人员并没有考虑到巡逻的响应时间，但是该因素已经在警察巡逻领域中被广泛研究[5-12]。结合响应时间来研究化工园区的巡逻问题十分有必要。

本文的研究方法本质上是博弈智能方法[13]，属于人工智能方法的一种。巡逻问题存在策略性的数据交互，传统数据驱动的手段无法预知在采用新的巡逻策略后对手的策略性响应，导致巡逻达不到预想的效果。博弈智能涵盖博弈论和人工智能等方向，重点研究个体或组织间的交互过程，以及如何通过构建博弈模型来实现最优策略的精确求解，最终形成智能化决策。

本文提出了面向化工园区有害气体泄漏的博弈巡逻策略。首先，运用安全博弈论的相关知识进行问题建模，考虑到有害气体泄漏的持续性，采用对流扩散模型[14]对扩散行为进行建模，同时将响应时间作为博弈模型的收益要素。然后，提出GGC算法对该模型进行求解，得到Stackelberg均衡解。最后，在3种不同规模的真实化工园区案例场景下，与其他两种基础方法进行收益对比。

1　相关工作

1.1　化工园区中的安防博弈

博弈论理论不断发展，在化工园区领域的应用有了新的进展。Reniers等人[15-18]提出的安防博弈模型根据“一企一档”数据系统地研究化工园区内化工厂的投资行为和安全协作模型，以此预防“外部多米诺效应”。Zhang等人[19]将博弈论引入定性和半定量的安全风险评估方法中，构建智能对手模型以提升对化工厂的保护，然后基于化工园区的地理信息数据和“一企一档”数据提出化工厂保护博弈模型[20]，该模型能有效避免攻击者的恐怖袭击。Zhu等人[21]提出了针对固定监测站的化工园区大气环保博弈模型，并结合源项估计算法，实现了对监测资源的调度和对排放源的检测。Feng等人[22]将博弈论和风险评估数据结合，优化了拥有多种化工设施的城市的防守资源配置。

1.2　巡逻问题

巡逻作为减少犯罪、预防安全问题的最有效、最直接的手段，在维护公共安全方面发挥着重要作用。巡逻问题包括巡逻区域划分与巡逻资源配置、巡逻路线规划这两方面的问题[23]。从20世纪70年代到现在，针对这两方面问题的研究方法，经历了从经典搜索论方法、蒙特卡洛方法、基于图论的启发式优化方法[24]到人工智能相关方法[25]的发展变化。

对于巡逻区域划分与巡逻资源配置问题，可以根据区域的地理位置和重要性，将巡逻区域划分成多个面积较小的巡逻片区，再对多个巡逻片区进行

合理的资源配置，也可以不划分巡逻片区，将巡逻资源分配到整个巡逻区域。Camacho-Collados 等人[26]开发了一种基于启发式局部搜索的分区算法，在2个巡逻片区的情况下将目标收益提高了11.97%，在6个巡逻片区的情况下提高了10.40%。Chen 等人[27]提出的多线性整数规划，考虑了目标函数中工作负载平均值和偏差的凸组合，并结合图分割与禁忌搜索提供解决方案。Meng 等人[28]基于图论方法，提出按照响应时间、经济效益、工作平衡和见警率4个标准来划分区域，解决交通巡警服务平台设置的问题。

巡逻路线规划问题的目的是降低犯罪率，快速响应调度信息，保护巡逻区域中各个目标点的安全。Rosenshine[29]和 Olson 等人[30]都利用蒙特卡洛方法来规划巡逻队的巡逻路线。Xie 等人[31]将巡逻路线规划问题看作双人博弈问题，提出一种多智能体强化学习算法来规划最佳巡逻路线。Agmon 等人[32]通过马尔可夫决策模型确定概率函数，规划后的巡逻路线具有对抗入侵者的能力。Buermann 等人[33]使用枚举组合表示4种入侵方式的入侵检测概率，与迭代更新马尔可夫转换矩阵相比，该方法能够显著缩短运行时间，还可扩展到更复杂的巡逻环境。Huang 等人[34]提出了一种分布式事件驱动的协作策略以实现自主巡逻的目标，能够有效减少巡逻的平均时间和最大空闲时间。Jana 等人[35]提出了一种具有状态表示的马尔可夫决策过程（Markov decision process，MDP）的公式，深度Q网络使用经过新型奖励函数训练的有限长度状态向量估计MDP。该策略具有较强的自适应能力，能够适应代理位置的变化、不均匀的节点访问频率要求、表示环境的图形结构的变化以及轨迹中的随机性。

2 问题描述

本文考虑了在化工园区巡逻场景下的应急响应情况，研究面向化工园区有害气体泄漏的博弈巡逻策略。本节分别介绍对流扩散模型、化工园区网络图建模、博弈巡逻模型。

2.1 对流扩散模型

对流扩散模型是由稳态对流扩散方程推导而来的，需要估计源项参数 $\theta_s=(r_s,Q_s)$。其中 r_s 表示泄漏源的位置坐标，Q_s 表示泄漏强度。此外，D 表示泄漏物质在空气中的扩散速率，V_ω 表示平均风速，由此得到任意位置 $r_k=(x_k,y_k)$ 的泄漏物质浓度（有害气体浓度）为：

$$C\left(r_k|\theta_s\right)=\frac{Q_s}{4\pi D|r_k-r_s|}\mathrm{e}^{\frac{-V_\omega(y_k-y_s)}{2D}}\mathrm{e}^{\frac{-|r_k-r_s|}{\lambda}} \tag{1}$$

其中，$|r_k-r_s|$ 表示当前检测位置与泄漏源位置的直线距离。$\lambda=\sqrt{\dfrac{D\tau}{1+\dfrac{V_\omega^2\tau}{4D}}}$，$\tau$ 表示有害气体分子的平均寿命。由于传感器在检测气体浓度时会产生较大的误差，引入斯莫鲁霍夫斯基公式，将任意位置检测到的有害气体浓度转换为传感器与有害气体分子在单位时间内的平均接触次数，表示为：

$$R\left(r_k|\theta_s\right)=4\pi D\cdot a\cdot C\left(r_k|\theta_s\right) \tag{2}$$

其中，a 表示球形气体传感器的半径。

接下来，引入泊松分布建模传感器与有害气体分子的接触过程，$r_k=(x_k,y_k)$ 位置处的传感器单位时间接触到 d_k 次有害气体分子的概率为：

$$P\left(d_k(r_k)|\theta_s\right)=\frac{R\left(r_k|\theta_s\right)^{d_k}}{d_k!}\mathrm{e}^{-R\left(r_k|\theta_s\right)} \tag{3}$$

2.2 化工园区网络图建模

化工园区网络图建模将化工园区转换为无向图 $G(V,E)$，其中 V 代表所有化工厂车辆出入口和道路交叉口的集合，E 代表所有道路集合（包括一个化工厂的多个车辆出入口之间的内部通道）。在网络图模型（如图1所示）中，类似“A1”的大写字母与数字的组合代表相应化工厂的车辆出入口，类似“Cr1”的“Cr”与数字的组合代表道路交叉口，也就是巡逻队的初始位置。在巡逻过程中，巡逻队每到达一个节点（化工厂车辆出入口或道路交叉口）v_i，就利用自身配备的球形气体传感器来检测当前位置单位时间接触到 d_k 次有害气体分子的概率。当接触概率 $P\left(d_k(r_k)|\theta_s\right)$ 大于阈值 P_{lower}（传感器能够进行有害气体溯源的最小接触概率）时，巡逻队可以利用 Entrotaxis 算法[36]找到泄漏化工厂的大概位置，及时制止事故。假定巡逻队在巡逻过程中始终保持匀速行驶，且园区道路上没有障碍物，同时巡逻队车辆始终处于有油或有电的状态。在园区网络上巡逻时，巡逻队在道路 e_i 上的行驶时间或转移时间可以由 t_{e_i} 表示。当巡逻时间 t_p 达到给定最大巡逻时长 T 时，巡逻结束。

在实际巡逻过程中，时间是一个连续变量，t_{e_i}

和t_p一般情况下不会是整数，并且泄漏事故可能会在巡逻持续时间[0,T]中的任意时刻发生。如果考虑时间的连续化，问题规模将会十分庞大而无法求解。为此，本文对巡逻模型的时间进行离散化处理，把巡逻持续时间[0,T]划分成多个相等的时间片，每个时间片都表示1分钟。所有与时间相关的非整数参数都需要进行离散化处理，四舍五入为整数。比如，巡逻队在道路e_i上的转移时间t_{e_i}为8.2分钟，t_{e_i}经过离散化处理后应该为8分钟。泄漏事故的发生时间只能在每个时间片的开始处，即单个时间片的整数倍。由此，可以列举出泄漏事故的所有发生情况。

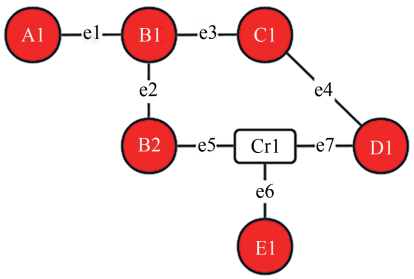


图1　化工园区网络图模型示例

2.3　博弈巡逻模型

在化工园区博弈巡逻中，参与者分别是以园区巡逻队为代表的“防守者”和以造成有害气体泄漏事故的不法分子为代表的“攻击者”。巡逻队在博弈中处于“领导者”位置，先行采取巡逻策略，而不法分子位于“跟随者”位置，通过一段时间对“领导者”活动的观察来获得巡逻策略，然后选择最有利于自身收益的攻击策略，造成目标化工厂发生有害气体泄漏事故，同时尽可能规避被巡逻队发现的风险。巡逻队希望选择最优巡逻策略，较大概率检测到空气中的有害气体分子，进一步实现对目标化工厂的溯源并及时制止泄漏事故。在博弈中，假设参与者都是完全理性的人。

攻击者的纯策略由造成有害气体泄漏的开始时间t_s^l、发生泄漏事故的目标化工厂p_l和规定响应时间rt_l这3部分组成。其中，规定响应时间也可以理解为有害气体泄漏的持续时间，但是为了在巡逻队的响应时间与规定响应时间进行收益计算，这里定义为规定响应时间更为合适。在$l\in[1,L]$中，L表示化工厂的数量，p_l表示第l个化工厂。攻击者针对目标化工厂p_l的纯策略$s_a^l\in S_a$可以表示为：

$$s_a^l=(t_s^l,p_l,rt_l) \tag{4}$$

其中，规定响应时间rt_l通常与发生泄漏事故的严重程度有关。为了简化问题，假定该参数是一个常数。

防守者的纯策略由多条转移边$((ct,cn),(nt,nd))$组合而成。其中，ct表示当前巡逻时间，cn表示当前巡逻队所处的位置节点，nt表示巡逻队到达下一个位置节点的巡逻时间，nd表示巡逻队的下一个位置节点。防守者的纯策略$s_{d_i}\in S_d$可以表示为：

$$s_{d_i}=\prod\nolimits_{((ct,cn),(nt,nd))\in pE}((ct,cn),(nt,nd)) \tag{5}$$

当化工厂p_l发生有害气体泄漏事故且检测概率$P(d_k(r_k)|\theta_s)$达到阈值时，则防守者的响应时间rt_d^l可以表示为：

$$rt_d^l=ct+cost(cn,p_l)-t_s^l \tag{6}$$

其中，$cost(cn,p_l)$表示从当前节点cn到目标化工厂p_l的最短时间。

最后，对防守者和攻击者的收益模型进行介绍。博弈巡逻的结果只有两种情况：攻击者成功和攻击者失败。针对目标化工厂p_l，如果攻击者成功，那么防守者面临损失L_d^l，攻击者获得收益G_a^l；如果攻击者失败，那么攻击者受到惩罚P_a^l，防守者得到奖励R_d^l。R_d^l是由化学园区管理部门决定的参数值，比如，管理部门奖励1 000元给防守者（由巡逻队和化工厂自身的安全部门组成）。P_a^l与巡逻场景相关，因为攻击不同的场景需要不同的攻击成本，如果攻击者被抓住，也将受到不同的惩罚。L_d^l和G_a^l是由相应成本和收益情况决定的。根据化工园区内的储罐数量和储罐中化工品在发生泄漏后的危害程度，化工领域内的专家提供L_d^l、G_a^l、P_a^l和R_d^l参数值，通过这些参数值得到博弈收益矩阵，它会影响博弈巡逻策略的规划。上述参数值代表着巡逻结果为防守者和攻击者带来的经济损失或收益，单位均为千元。

下面给出L_d^l、G_a^l、P_a^l和R_d^l的具体定义：

$$L_d^l=loss_d^l+cost_d \tag{7}$$

$$G_a^l=reward_a^l-cost_a \tag{8}$$

$$P_a^l=loss_a^l+cost_a \tag{9}$$

$$R_d^l=reward_d^l \tag{10}$$

其中，$loss_d^l$表示化工厂p_l在被攻击成功后，防守者面临的经济损失；$cost_d$表示防守者的防守成本；

$reward_a^l$表示化工厂p_l在被攻击成功后，攻击者获得的经济收益；$cost_a$表示攻击者的攻击成本；$loss_a^l$表示化工厂p_l在被攻击失败后，攻击者需要承担的经济罚款；$reward_d^l$表示化工厂p_l在被攻击失败后，防守者得到的经济奖励。本文考虑的是破坏化工厂并导致有害气体泄漏的攻击场景，$reward_a^l$和$reward_d^l$都是针对该攻击场景的经济收益。

下面给出防守者收益u_d^l和攻击者收益u_a^l的计算式：

$$u_d^l=R_d^l\cdot M-L_d^l\cdot N \tag{11}$$

$$u_a^l=G_a^l\cdot N-P_a^l\cdot M \tag{12}$$

其中，M代表巡逻队在当前位置节点的有害气体检测情况，N代表巡逻队在发生有害气体泄漏后的响应情况，具体表示为：

$$M=\begin{cases}1, P\left(d_{cn}\left(r_{cn}\right)|\theta\right)>P_{\text{lower}},\\ \quad \forall\left((ct,cn),(nt,nd)\right)\in s_d,\ ct\geqslant t_s^l\\ 0,\ \text{others}\end{cases} \tag{13}$$

$$N=\begin{cases}K\cdot rt_d^l,\ M=1, rt_d^l\leqslant rt_l\\ 1,\ \text{others}\end{cases} \tag{14}$$

其中，P_{lower}表示传感器能够进行有害气体溯源的最小接触概率，K表示归一化常数。

表1是博弈收益矩阵示例，行代表攻击者的两种攻击策略，分别是s_a^1和s_a^2；列代表防守者的两种防守策略，分别是s_{d_1}和s_{d_2}。在该收益矩阵中的值由元组表示，比如$(R_d^1\cdot M_{s_{d_1},s_a^1}-L_d^1\cdot N_{s_{d_1},s_a^1},G_a^1\cdot N_{s_{d_1},s_a^1}-P_a^1\cdot M_{s_{d_1},s_a^1})$表示在攻击者选择策略$s_a^1$和防守者选择策略$s_{d_1}$时，双方的收益情况，防守者的收益情况为$R_d^1\cdot M_{s_{d_1},s_a^1}-L_d^1\cdot N_{s_{d_1},s_a^1}$，攻击者的收益情况为$G_a^2\cdot N_{s_{d_1},s_a^1}-P_a^1\cdot M_{s_{d_1},s_a^1}$，$M_{s_{d_1},s_a^1}$代表在双方策略下$M$的取值情况，$N_{s_{d_1},s_a^1}$代表在双方策略下$N$的取值情况。

表1 博弈收益矩阵示例

策略	s_a^1	s_a^2
s_{d_1}	$(R_d^1\cdot M_{s_{d_1},s_a^1}-L_d^1\cdot N_{s_{d_1},s_a^1},G_a^1\cdot N_{s_{d_1},s_a^1}-P_a^1\cdot M_{s_{d_1},s_a^1})$	$(R_d^2\cdot M_{s_{d_1},s_a^2}-L_d^2\cdot N_{s_{d_1},s_a^2},G_a^2\cdot N_{s_{d_1},s_a^2}-P_a^2\cdot M_{s_{d_1},s_a^2})$
s_{d_2}	$(R_d^1\cdot M_{s_{d_2},s_a^1}-L_d^1\cdot N_{s_{d_2},s_a^1},G_a^1\cdot N_{s_{d_2},s_a^1}-P_a^1\cdot M_{s_{d_2},s_a^1})$	$(R_d^2\cdot M_{s_{d_2},s_a^2}-L_d^2\cdot N_{s_{d_2},s_a^2},G_a^2\cdot N_{s_{d_2},s_a^2}-P_a^2\cdot M_{s_{d_2},s_a^2})$

3 巡逻算法

本节将介绍面向化工园区有害气体泄漏的博弈巡逻模型的相关算法。首先，介绍化工园区巡逻转移图生成算法；然后，提出用于求解博弈巡逻模型的GGC算法。

3.1 巡逻转移图生成算法

为了更全面、直观地展现巡逻策略，构建巡逻转移图[2]是一种很好的表示方法。本节根据化工园区网络图及其邻接矩阵，在给定最大巡逻时长T的情况下，给出所有可能的巡逻策略，方便对不同巡逻策略进行分析与比较。巡逻转移图生成算法如算法1所示。

算法1：巡逻转移图生成算法

输入：网络图$G(V,E)$，邻接矩阵$cG(cn,nd)$，巡逻队初始状态$pv=(0,sn)$，给定最大巡逻时长T，数组tNL、巡逻队状态数组pV和巡逻队转移边数组pE

输出：巡逻队状态数组pV，巡逻队转移边数组pE

1 初始化 $tNL\leftarrow pv$, $pV\leftarrow pv$

2 while $tNL\neq\varnothing$ do

3 $cv=(ct,cn)\leftarrow tNL$ // 取出数组tNL的第一个元素并定义为当前状态$cv=(ct,cn)$

4 $ccn=\{nd\in V|cG(cn,nd)<\infty\}$//根据邻接矩阵$cG$找到巡逻队当前所处节点的所有邻居节点集合$ccn$

5 for each $nd\in ccn$ do

6 if $ct+cG(cn,nd)\leqslant T+cost(sn,nd)$ do

7 $nv=(ct+cG(cn,nd),nd)$//得到新状态nv

8 $ne=(cv,nv)$// 从当前状态cv转移到新状态nv的转移边ne

9 $pE\leftarrow ne$ // 将转移边ne加入转移边数组pE

10 if nv not in pV do

11 $tNL\leftarrow nv$ // 将新状态nv加入数组tNL

12 $pV\leftarrow nv$ // 将新状态nv加入状态数组pV

13 remove $cv$from tNL // 将当前状态cv从数组tNL中移除

14 end while

巡逻队在每一时刻的状态被定义为状态(t,i)，表示巡逻队在时间t位于节点i位置处的状态。从状态(t_m,i_m)到状态(t_n,i_n)的转移边表示位于节点i_m的巡逻队在时间t_m做出前往节点i_n并在时间t_n到达的决策。

图2是由图1得到的化工园区巡逻转移图示例，给出了在给定最大巡逻时长T=20内的所有巡逻策略，并用不同颜色的线段组合表示不同的巡逻策略。若巡逻时间超过给定的最大巡逻时长，但未到达巡逻目标节点，巡逻队仍会继续巡逻，因此实际巡逻时长会超过给定最大巡逻时长。黑色加粗实线表示一条完整的巡逻路线“Cr1—E1—Cr1—D1—Cr1—

D1—Cr1—E1—Cr1—B2—B1”。红色加粗实线表示当巡逻队采用黑色加粗实线作为固定巡逻策略时，攻击者的最优响应策略：在时间0时造成化工厂节点“E1”发生有害气体泄漏，规定响应时间为5分钟。

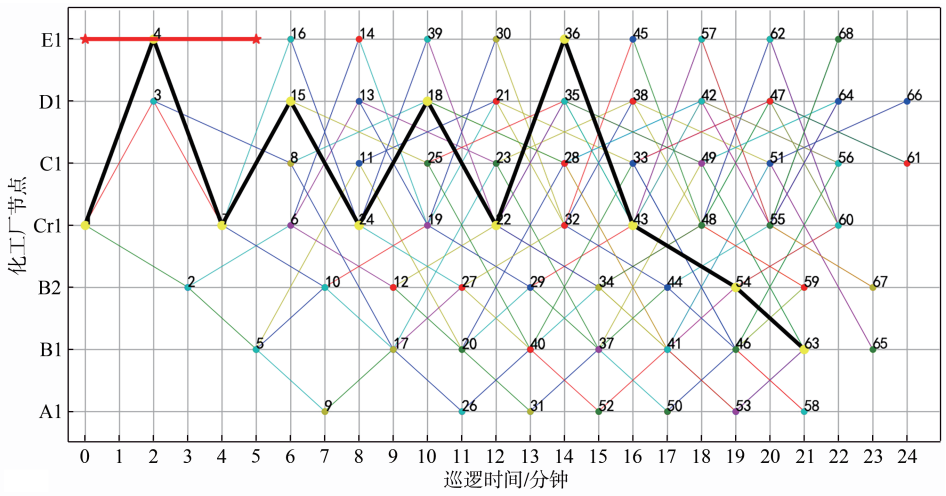


图2　化工园区巡逻转移图示例

3.2　GGC算法

本文提出的GGC算法是用于求解化工园区内有害气体泄漏的博弈巡逻策略算法，该算法是基于多线性规划算法实现的。多线性规划算法是由Conitzer和Sandholm[37]提出的、用于求解贝叶斯Stackelberg博弈中领导者最优策略的基准算法。该算法利用Stackelberg均衡定义，将求解领导者最优混合策略问题转化为优化问题。本文使用GGC算法得到攻防博弈双方的均衡策略：

$$\max_{s_{d_i}\in S_d}\ (s_{d_i},(t_s^l,p_l,rt_l)^*) \tag{15}$$

$$\text{maxPayoff}_d(s_{d_i}{}^*,(t_s^l,p_l,rt_l)) \tag{16}$$

$$u_a^l((t_s^l,p_l,rt_l)^*,s_{d_i})\geqslant u_a^l((t_s^l,p_l,rt_l),s_{d_i}),\forall\left(t_s^l,p_l,rt_l\right),\quad s_{d_i}\in S_d \tag{17}$$

在算法求解过程中，对于攻击者每一种纯策略$\left(t_s^l,p_l,rt_l\right)$，遍历防守者的所有纯策略并根据收益式（11）和式（12）计算得到u_d^l和u_a^l。式（15）是防守者的优化目标，其中$(t_s^l,p_l,rt_l)^*$是攻击者的最佳响应策略。式（16）是攻击者的优化目标，其中$s_{d_i}{}^*$是防守者的最优策略。式（17）是为了从攻击者的纯策略集合中找到最优响应策略而设定的约束条件。与此同时，防守者选择每个纯策略的概率$Prob(s_{d_i})$应满足以下条件：

$$\sum\nolimits_{s_{d_i}\in S_d} Prob(s_{d_i})=1 \tag{18}$$

对攻击者的每种纯策略进行线性规划求解，可获得博弈巡逻下的Stackelberg均衡解，最终防守者得到最优混合巡逻策略。

4　实验分析

本文在3种不同规模的真实化工园区案例场景下，对面向化工园区有害气体泄漏的博弈巡逻模型进行实验。

4.1　实验设置

本文分别在中国的钦州化工园区、中国的上海化工园区和比利时的Antwerp化工园区的场景下进行了实验。假设在上述化工园区案例中有害气体泄漏模型的参数均为：$D=50, V_\omega=5, \tau=150, a=5, d_k=10$。巡逻队从初始位置“Cr1”开始巡逻任务。巡逻实验的规定响应时间由化工园区内的道路规模大小决定，具体参数由园区管理部门设定。实验收益矩阵由化工领域内的专家提供。这3种化工园区的道路网络图参数如表2所示。

表2　化工园区道路网络图参数

化工园区	节点数	连边数	化工厂节点数	道路交叉口数
钦州化工园区	7	7	5	2
上海化工园区	10	15	7	3
Antwerp化工园区	16	25	9	7

4.2　钦州化工园区案例

钦州化工园区是一个小规模的化工园区，共有7个节点，分别是“A1”、“B1”、“C1”、

“D1”、“E1”、“Cr1”和“Cr2”。园区布局及网络图模型如图 3 所示。博弈模型的收益矩阵如表 3 所示。其中，针对目标化工厂 p_l，如果攻击者成功，那么防守者面临损失 L_d^l，攻击者获得收益 G_a^l；如果攻击者失败，那么攻击者受到惩罚 P_a^l，防守者得到奖励 R_d^l。这些参数值表示巡逻结果给防守者和攻击者带来的经济损失或收益，单位为千元。

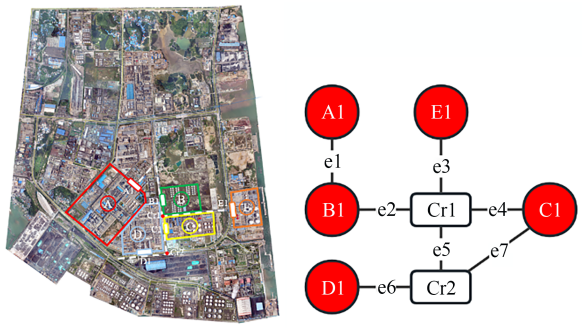


图 3　钦州化工园区布局（左）及其网络图模型（右）

表 3　钦州化工园区案例场景的收益矩阵（单位：千元）

化工厂	R_d^l	L_d^l	G_a^l	P_a^l
A	2	15	10	4
B	2	10	6.5	4
C	2	13	8	4
D	2	12	7	4
E	2	16	9	4

图 4 展示了在钦州化工园区案例场景下进行博弈巡逻得到的 Stackelberg 均衡。其中，黑色加粗实线表示巡逻队的最佳巡逻策略。在该均衡解中，攻击者的最佳策略是在时间 3 时对化工厂节点“A1”发起攻击并造成有害气体泄漏，规定响应时间为 3 分钟，由红色加粗实线表示。在该博弈均衡下，根据式（15）~式（17），计算得出防守者的最佳收益为-6.55，攻击者的最佳收益为 2.38。巡逻队的最优混合策略如表 4 所示。攻击者的最优策略如表 5 所示，每条策略对于攻击者来说都是最优的。

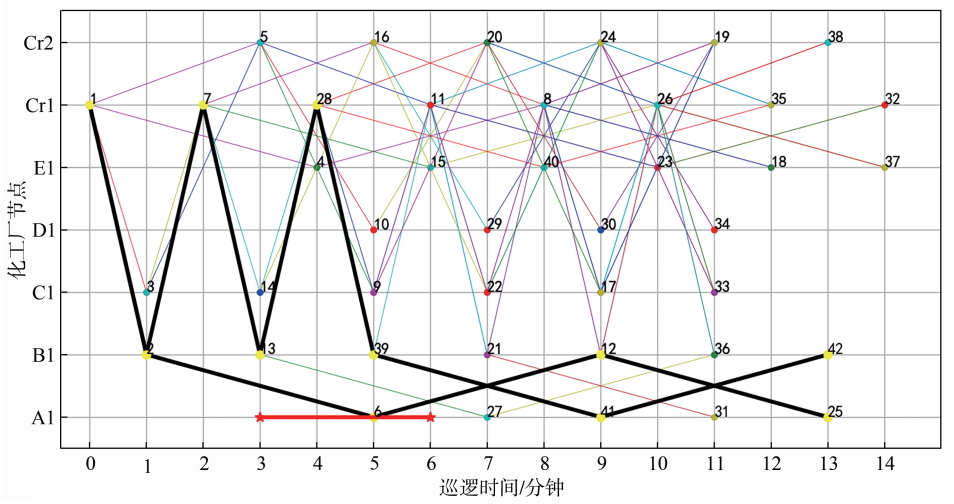


图 4　钦州化工园区案例场景下的均衡解

表 4　巡逻队的最优混合策略

策略编号	巡逻策略	选择概率
0	Cr1-B1-A1-B1-A1	0.5
1	Cr1-B1-Cr1-B1-Cr1-B1-A1-B1	0.5

表 5　攻击者的最优策略

策略编号	攻击策略
0	（A1,3,3）
1	（A1,4,3）
2	（A1,5,3）
3	（A1,7,3）

4.3　上海化工园区案例

上海化工园区是一个中等规模的化工园区，其中共有 10 个节点，分别是“A1”、“B1”、“B2”、“B3”、“C1”、“C2”、“D1”、“Cr1”、“Cr2”和“Cr3”。园区布局如图 5 所示，网络图模型如图 6 所示。博弈模型的收益矩阵如表 6 所示。

图 7 展示了在上海化工园区案例场景下博弈巡逻得到的 Stackelberg 均衡。黑色加粗实线表示巡逻队的最佳巡逻策略。在该均衡解中，攻击者的最佳策略是在时间 0 时对化工厂节点“B2”发起

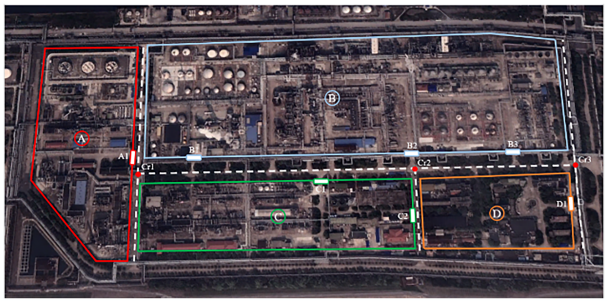


图5　上海化工园区布局

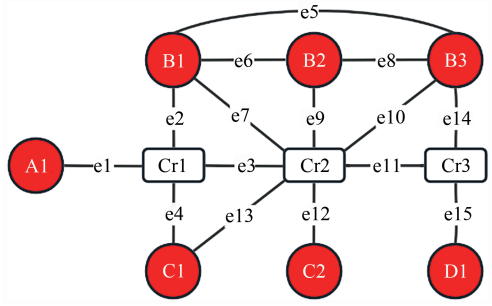


图6　上海化工园区网络图模型

表6　上海化工园区案例场景的收益矩阵（单位：千元）

化工厂	R_d^l	L_d^l	G_a^l	P_a^l
A	2	10	10	4
B	2	14	12	4
C	2	11	8	4
D	2	9	6	4

攻击并造成有害气体泄漏，规定的响应时间为10分钟，由红色加粗实线表示。该博弈均衡下，根据式（15）~式（17），计算得出防守者的最佳收益为-8.27，攻击者的最佳收益为3.13。巡逻队的最优混合策略如表7所示。攻击者的最优策略如表8所示，其中每条策略对于攻击者来说都是最优的。

4.4　Antwerp化工园区案例

Antwerp化工园区是一个大规模的化工园区，共有16个节点，分别是A1、A2、B1、C1、D1、E1、F1、G1、H1、Cr1、Cr2、Cr3、Cr4、Cr5、Cr6和Cr7。园区布局如图8所示，网络图模型如图9所示。博弈模型的收益矩阵如表9所示。

图10展示了在Antwerp化工园区案例场景下博弈巡逻得到的Stackelberg均衡。黑色加粗实线表示巡逻队的最佳巡逻策略。在该均衡解中，攻击者的最佳策略是在时间1时对化工厂节点“A1”发起攻击并造成有害气体泄漏，规定的响应时间为10分钟，由红色加粗实线表示。在该博弈均衡下，根据式（15）~式（17），计算得出防守者的最佳收益为-9.34，攻击者的最佳收益为4.95。巡逻队的最优混合策略如表10所示。攻击者的最优策略如表11所示，每条策略对于攻击者来说都是最优的。

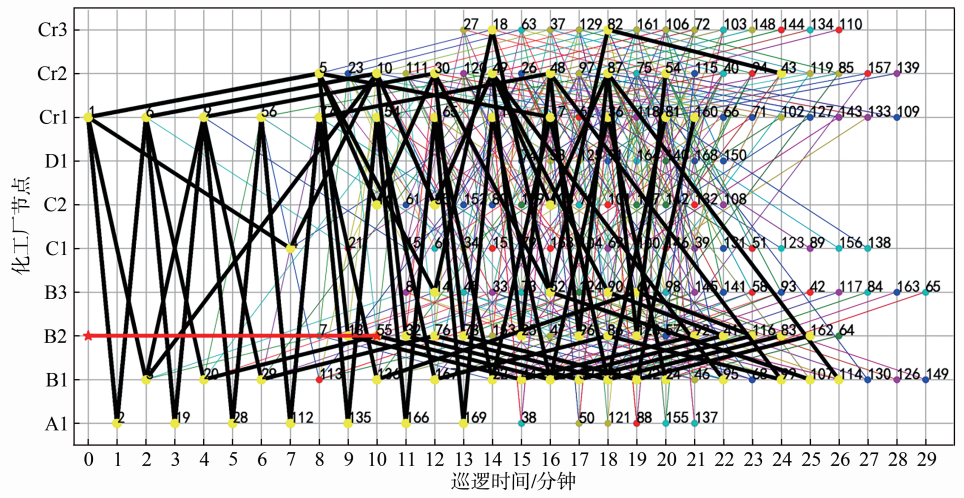


图7　上海化工园区案例场景下的均衡解

4.5　收益结果对比

采用随机巡逻方法和基于GGP（global greedy patrolling）算法的巡逻方法[38]规划巡逻队的巡逻策略，并将得到的巡逻策略作为博弈巡逻模型中防守者选取的策略，攻击者对该策略做出最佳响应，计算博弈双方各自的收益结果。不同方法在不同化工园区案例中的收益结果如表12所示。

随着巡逻算法的优化，防守者的收益逐渐提

表 7　巡逻队的最优混合策略

策略编号	巡逻策略	选择概率
0	Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-B1-Cr1-B1	0.0163
1	Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-B1-Cr1-B1-B2	0.0599
2	Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-B1-B2-B1	0.1299
3	Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-B1-B3-B2	0.0058
4	Cr1-A1-Cr1-A1-Cr1-A1-Cr1-A1-Cr1-Cr2-B2-Cr2-B1	0.0380
5	Cr1-A1-Cr1-A1-Cr1-Cr2-B2-B1-B2	0.0271
6	Cr1-A1-Cr1-B1-B2-B1-B2	0.0459
7	Cr1-A1-Cr1-Cr2-B2-B1-B2	0.0545
8	Cr1-A1-Cr1-Cr2-C2-Cr2-B3-B2	0.0172
9	Cr1-B1-Cr1-B1-B2-B1-B2	0.0260
10	Cr1-B1-Cr2-C2-Cr2-B1	0.1086
11	Cr1-B1-Cr2-C2-Cr2-C2-Cr2-B2-Cr2	0.0630
12	Cr1-C1-Cr2-C2-Cr2-B2-Cr2-B1	0.0358
13	Cr1-Cr2-B2-B1-B2	0.1433
14	Cr1-Cr2-B3-Cr3-B3-B1	0.1433
15	Cr1-Cr2-C2-Cr2-B2-B1-Cr1	0.0146
16	Cr1-Cr2-C2-Cr2-B3-Cr3-Cr2	0.0372
17	Cr1-Cr2-Cr1-B1-Cr1	0.0334

表 8　攻击者的最优策略

策略编号	攻击策略
0	（B2，0，10）
1	（B3，3，10）
2	（B3，4，10）
3	（B3，5，10）

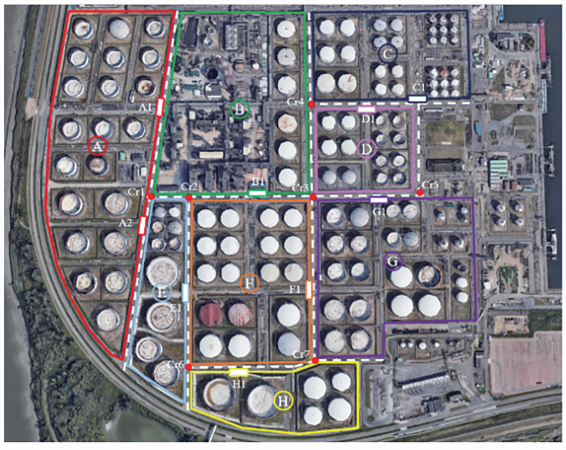


图 8　Antwerp 化工园区布局

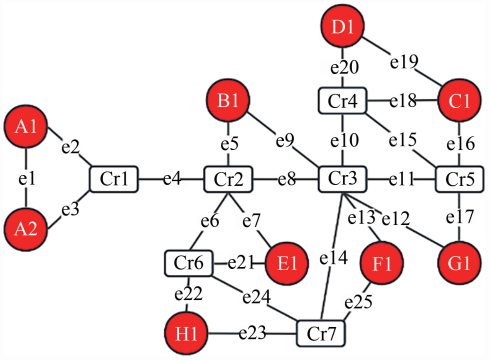


图 9　Antwerp 化工园区网络图模型

表 9　Antwerp 化工园区案例场景的收益矩阵（单位：千元）

化工厂	R_d^l	L_d^l	G_a^l	P_a^l
A	3	18	15	5
B	3	15	6	5
C	3	10	10	5
D	3	8	9	5
E	3	7	8	5
F	3	12	11	5
G	3	13	10	5
H	3	7	7	5

高，而攻击者的收益不断下降。由于随机方法的随机性，防守者的收益是最差的，而攻击者的收益是最好的，这无疑会生成较差的巡逻策略。基于 GGP 的方法，利用贪心搜索来规划巡逻策略，其目标是最小化全局平均响应时间，从而提高博弈收益。基于 GGP 的方法得到的巡逻策略是近似解，而本文方法遍历攻防双方的所有纯策略，再利用 GGC 算法得到最优巡逻策略，得到的策略是精确

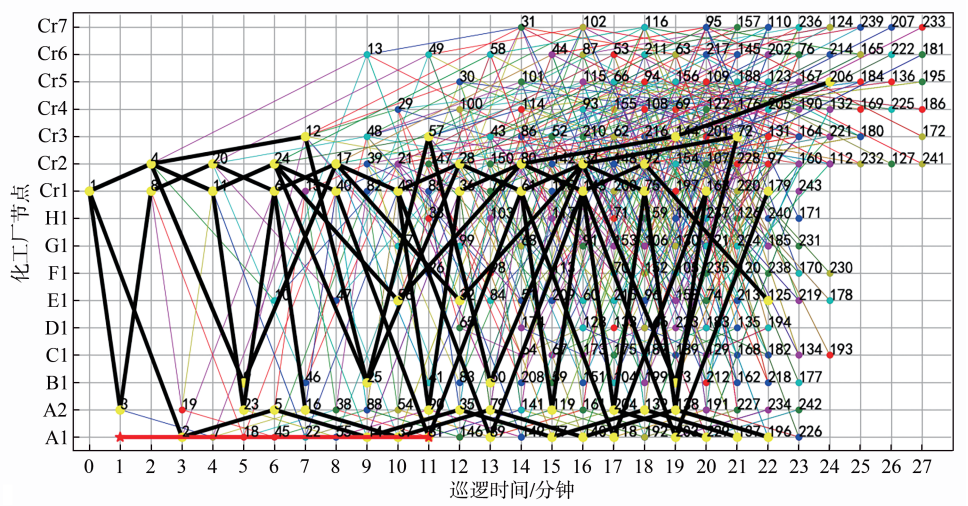


图 10　Antwerp 化工园区案例场景下的均衡解

表 10　巡逻队的最优混合策略

策略编号	巡逻策略	选择概率
0	Cr1-A1-A2-A1-A2-A1-A2-A1	0.2018
1	Cr1-A2-Cr1-Cr2-Cr1-A2-A1-A2-A1-A2-A1	0.2018
2	Cr1-Cr2-B1-Cr2-Cr1-A1-Cr1-Cr2-E1	0.0242
3	Cr1-Cr2-B1-Cr2-Cr1-A2-Cr1-Cr2-Cr3-Cr5	0.0242
4	Cr1-Cr2-B1-Cr2-Cr1-Cr2-Cr1-A2-Cr1-A2-A1	0.1251
5	Cr1-Cr2-Cr1-A2-Cr1-Cr2-E1-Cr2-Cr3	0.0767
6	Cr1-Cr2-Cr1-Cr2-B1-Cr2-Cr1-A1-Cr1	0.1209
7	Cr1-Cr2-Cr1-Cr2-E1-Cr2-Cr1-A1-Cr1	0.0155
8	Cr1-Cr2-Cr1-Cr2-E1-Cr2-Cr1-A2-Cr1-A2-A1	0.0913
9	Cr1-Cr2-Cr1-Cr2-Cr1-A1-Cr1-Cr2-B1-Cr3	0.0767
10	Cr1-Cr2-Cr3-B1-Cr3-B1-Cr2-Cr1-A2-Cr1	0.0417

表 11　攻击者的最优策略

策略编号	攻击策略
0	（A1，0，10）
1	（A1，1，10）

表 12　不同方法在不同化工园区案例中的收益结果

园区名	随机方法	基于 GGP 的方法	博弈方法
钦州化工园区	（-8.42，3.52）	（-6.86，2.78）	（-6.55，2.38）
上海化工园区	（-11.07，4.87）	（-9.40，3.68）	（-8.27，3.13）
Antwerp 化工园区	（-13.53，6.11）	（-9.81，5.09）	（-9.34，4.95）

解，收益表现更好。上述收益结果表明，本文方法能够有效提升防守者的收益，降低攻击者的收益，从而降低攻击者对化工厂发动攻击的概率。

5　结束语

针对目前化工园区内巡逻策略效果有限的问题，本文采用博弈智能方法，提出面向化工园区有害气体泄漏的博弈巡逻模型，详细地给出博弈模型中参与者、策略、收益和均衡解的建模方法，有害气体扩散模型的加入使得巡逻场景更加真实。在收益模型中的响应时间约束，要求巡逻队不仅要及时发现化工事故，还要在规定时间内到达事故处，从而获得尽可能大的收益。此外，将该博弈巡逻模型与其他两种基础算法的收益结果进行比较，结果表明该博弈巡逻模型是有效的。本文提出的博弈巡逻模型对于解决化工行业的安全问题具有借鉴意义。

参考文献：

[1] 赵刚. 城市人民警察巡逻机制建设二十年回眸与展望[J]. 公安研究, 2014(2): 25-32.
ZHAO G. Twenty years review and prospect of urban people's police patrol mechanism[J]. Policing Studies, 2014(2): 25-32.

[2] ZHANG L B, RENIERS G, CHEN B, et al. CCP game: a game theoretical model for improving the scheduling of chemical cluster patrolling[J]. Reliability Engineering & System Safety, 2019, 191: 106186.

[3] REZAZADEH A, ZHANG L B, RENIERS G, et al. Optimal patrol scheduling of hazardous pipelines using game theory[J]. Process Safety and Environmental Protection, 2017, 109: 242-256.

[4] 陈斐然, 陈彬, 朱正秋, 等. 基于博弈论的化工园区合作巡逻策略研究[J]. 系统仿真学报, 2020, 32(10): 1903-1909.
CHEN F R, CHEN B, ZHU Z Q, et al. Chemical cluster cooperative patrolling strategy based on game theory[J]. Journal of System Simulation, 2020, 32(10): 1903-1909.

[5] GHOSH S, VARAKANTHAM P. Dispatch guided allocation optimization for effective emergency response[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1).

[6] LEI C, LIN W H, MIAO L X. A stochastic emergency vehicle redeployment model for an effective response to traffic incidents[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 16(2): 898-909.

[7] LARSON R C. Urban police patrol analysis[M]. Cambridge: MIT

Press, 1972.
[8] LIBERATORE F, CAMACHO-COLLADOS M. A comparison of local search methods for the multicriteria police districting problem on graph[J]. Mathematical Problems in Engineering, 2016, 2016: 1-13.
[9] KUO P F, LORD D, WALDEN T D. Using geographical information systems to organize police patrol routes effectively by grouping hotspots of crash and crime data[J]. Journal of Transport Geography, 2013, 30: 138-148.
[10] COUPE R T, GIRLING A J. Modelling police success in catching burglars in the act[J]. Omega, 2001, 29(1): 19-27.
[11] CAMACHO-COLLADOS M, LIBERATORE F. A decision support system for predictive police patrolling[J]. Decision Support Systems, 2015, 75: 25-37.
[12] D'AMICO S J, WANG S J, BATTA R, et al. A simulated annealing approach to police district design[J]. Computers & Operations Research, 2002, 29(6): 667-684.
[13] 郝建业, 邵坤, 李凯, 等. 博弈智能的研究与应用[J]. 中国科学: 信息科学, 已录用,doi: 10.1360/SSI-2023-0010.
HAO J Y, SHAO K, LI K, et al. Research and applications of game intelligence[J]. Scientia Sinica(Informationis), accepted, doi: 10.1360/SSI-2023-0010.
[14] VERGASSOLA M, VILLERMAUX E, SHRAIMAN B I. 'Infotaxis' as a strategy for searching without gradients[J]. Nature, 2007, 445(7126): 406-409.
[15] RENIERS G, DULLAERT W, KAREL S. Domino effects within a chemical cluster: a game-theoretical modeling approach by using Nash-equilibrium[J]. Journal of Hazardous Materials, 2009, 167(1-3): 289-293.
[16] RENIERS G, SOUDAN K. A game-theoretical approach for reciprocal security-related prevention investment decisions[J]. Reliability Engineering & System Safety, 2010, 95(1): 1-9.
[17] RENIERS G. An external domino effects investment approach to improve cross-plant safety within chemical clusters[J]. Journal of Hazardous Materials, 2010, 177(1-3): 167-174.
[18] RENIERS G, CUYPERS S, PAVLOVA Y. A game-theory based multi-plant collaboration model (MCM) for cross-plant prevention in a chemical cluster[J]. Journal of Hazardous Materials, 2012, 209-210: 164-176.
[19] ZHANG L B, RENIERS G, CHEN B, et al. Integrating the API SRA methodology and game theory for improving chemical plant protection [J]. Journal of Loss Prevention in the Process Industries, 2018, 51: 8-16.
[20] ZHANG L B, RENIERS G. A game-theoretical model to improve process plant protection from terrorist attacks[J]. Risk Analysis, 2016, 36 (12): 2285-2297.
[21] ZHU Z Q, CHEN B, RENIERS G, et al. Playing chemical plant environmental protection games with historical monitoring data[J]. International Journal of Environmental Research and Public Health, 2017, 14 (10): 1155.
[22] FENG Q L, CAI H, CHEN Z L, et al. Using game theory to optimize allocation of defensive resources to protect multiple chemical facilities in a city against terrorist attacks[J]. Journal of Loss Prevention in the Process Industries, 2016, 43: 614-628.
[23] SAMANTA S, SEN G, GHOSH S K. A literature review on police patrolling problems[J]. Annals of Operations Research, 2022, 316(2): 1063-1106.
[24] DANTZIG G, FULKERSON R, JOHNSON S. Solution of a large-scale traveling-salesman problem[J]. Journal of the Operations Research Society of America, 1954, 2(4): 393-410.
[25] LI W Q, CHEN D H, LE J J. Robot patrol path planning based on combined deep reinforcement learning[C]//Proceedings of 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications. Piscataway: IEEE Press, 2018: 659-666.
[26] CAMACHO-COLLADOS M, LIBERATORE F, ANGULO J M. A multi-criteria police districting problem for the efficient and effective design of patrol sector[J]. European Journal of Operational Research, 2015, 246(2): 674-684.
[27] CHEN H F, CHENG T, YE X Y. Designing efficient and balanced police patrol districts on an urban street network[J]. International Journal of Geographical Information Science, 2019, 33(2): 269-290.
[28] MENG J X, LU X Z. The application of graph theory method in the problem of traffic patrol police service platform's site selection[J]. Applied Mechanics and Materials, 2012, 246/247: 723-727.
[29] ROSENSHINE M. Contributions to a theory of patrol scheduling[J]. Journal of the Operational Research Society, 1970, 21(1): 99-106.
[30] OLSON D G, WRIGHT G P. Models for allocating police preventive patrol effort[J]. Journal of the Operational Research Society, 1975, 26 (4): 703-715.
[31] XIE Y, GU Y J. Police patrol path planning using stackelberg equilibrium based multiagent reinforcement learning[J]. Transactions of Beijing Institute of Technology, 2017, 37(1): 93-99.
[32] LIN E S, AGMON N, KRAUS S. Multi-robot adversarial patrolling: handling sequential attacks[J]. Artificial Intelligence, 2019, 274: 1-25.
[33] BUERMANN J, ZHANG J. Multi-robot adversarial patrolling strategies via lattice paths[J]. Artificial Intelligence, 2022, 311: 103769.
[34] HUANG L, ZHOU M C, HAO K R, et al. Multirobot cooperative patrolling strategy for moving objects[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 53(5): 2995-3007.
[35] JANA M, VACHHANI L, SINHA A. A deep reinforcement learning approach for multi-agent mobile robot patrolling[J]. International Journal of Intelligent Robotics and Applications, 2022, 6(4): 724-745.
[36] HUTCHINSON M, OH H, CHEN W H. Entrotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions[J]. Information Fusion, 2018, 42: 179-189.
[37] CONITZER V, SANDHOLM T. Computing the optimal strategy to commit to[C]// Proceedings of the 7th ACM Conference on Electronic Commerce. New York: ACM, 2006: 82-90.
[38] ZHANG Y P, XIAO Y, WANG Y, et al. Bio-inspired patrolling scheme design in wireless and mobile sensor and robot networks[J]. Wireless Personal Communications, 2017, 92: 1303-1332.

[作者简介]



陈茵（1983– ），女，就职于中国电信股份有限公司广西分公司，主要研究方向为大数据分析、信息化项目管理。



张立泽（1997– ），男，杭州电子科技大学硕士生，主要研究方向为安防巡逻。



陈黎丽（1985– ），女，博士，杭州电子科技大学讲师，主要研究方向为网络安全、博弈论、数据安全。



帅国华（1976– ），男，中国广西自贸区钦州港片区应急局注册安全工程师、局长，主要研究方向为化工园区安全生产监督管理。



王震（1984– ），男，博士，杭州电子科技大学教授，卓越学院副院长，主要研究方向为智能博弈对抗、人工智能安全、博弈论、多智能体系统、网络科学和去中心化安全。