

doi: 10.3969/j.issn.1001-893X.2021.11.007

引用格式: 陈泽盛, 杨承志, 曹鹏宇, 等. 一种基于双 DQN 的空战干扰样式选择方法[J]. 电讯技术, 2021, 61(11): 1371-1377. [CHEN Zesheng, YANG Chengzhi, CAO Pengyu, et al. An air combat jamming style selection method based on double DQN[J]. Telecommunication Engineering, 2021, 61(11): 1371-1377.]

一种基于双 DQN 的空战干扰样式选择方法^{*}

陈泽盛¹, 杨承志¹, 曹鹏宇¹, 邴雨晨¹, 纳 贤²

(1. 空军航空大学, 长春 130022; 2. 中国人民解放军 95510 部队, 贵阳 550031)

摘 要: 针对深度 Q 网络(Deep Q Network, DQN)的干扰样式选择方法, 提出了一种针对空战场景的双深度 Q 网络(Double Deep Q Network, Double DQN)干扰样式选择方法。以机载火控雷达空-空模式下的典型工作模式和我方干扰样式集为基础建立干扰矩阵作为算法输入, 结合 Double DQN 原理提出了干扰样式选择方法。通过仿真实验对比了 Double DQN 和 DQN 算法的性能, 并对部分参数设置进行探索, 给出了两种空战态势下的最优干扰样式选择路径。仿真结果表明, 该方法能够在满足实时性的条件下根据干扰效果在空战态势下选择出最优干扰样式。

关键词: 机载火控雷达; 干扰样式选择; 双 DQN; 干扰矩阵

开放科学(资源服务)标识码(OSID):



微信扫描二维码
听独家语音释文
与作者在线交流
享本刊专属服务

中图分类号: TN974 文献标志码: A 文章编号: 1001-893X(2021)11-1371-07

An Air Combat Jamming Style Selection Method Based on Double DQN

CHEN Zesheng¹, YANG Chengzhi¹, CAO Pengyu¹, BING Yuchen¹, NA Xian²

(1. Aviation University of Air Force, Changchun 130022, China; 2. Unit 95510 of PLA, Guiyang 550031, China)

Abstract: For the jamming style selection method of deep Q network (DQN), a Double Deep Q Network (Double DQN) jamming style selection method for air combat scene is proposed. Based on the typical working mode of the airborne fire control radar in air-to-air mode and our jamming style set, the jamming matrix is established as the algorithm input. Then a method for selecting jamming styles is proposed according to the principle of Double DQN. Finally, the performance of Double DQN algorithm and DQN algorithm is compared through simulation experiments, some parameter settings are explored, and the optimal jamming style selection path under two air combat situations is given. The results show that the method can select the optimal jamming style in air combat situations according to the jamming effect under real-time conditions.

Key words: airborne fire control radar; jamming style selection; double deep Q network; jamming matrix

0 引 言

随着机载火控雷达(Airborne Fire Control Radar, AFRCR)技术体制的不断发展, 其工作状态更

加复杂, 波形单元更灵活多变, 抗干扰能力更强^[1], 传统的基于人工加载的干扰样式选择方法^[2]将失去用武之地。在加载干扰样式正确的情况下干扰机

* 收稿日期: 2021-03-14; 修回日期: 2021-04-07
基金项目: 国防科技卓越青年基金资助项目(315090303)
通信作者: 陈泽盛

可以产生干扰效果,一旦敌方采取抗干扰措施,我方缺乏先验知识将无法选择合适的干扰样式,作战效能会大打折扣。强化学习由于无需先验知识,且其与环境交互的过程与雷达对抗相似,因此更适用于雷达干扰样式选择。

Q-学习是强化学习中的一种高效的免模型算法。文献[3]设计了基于Q-学习的认知雷达对抗过程,使雷达干扰系统通过自主学习实现高效干扰,但环境雷达的工作状态较少。文献[4]提出了智能雷达对抗(Intelligent Radar Countermeasure,IRC)方法,在工作模式数目未知的情况下通过Q-学习选择最优的干扰路径。文献[5]研究了对多功能雷达(Multifunctional Radar,MFR)的Q-学习认知干扰决策方法,对仿真中各参数对实验的影响进行分析,对比了该方法与其他传统方法的优势。但基于Q-学习的认知干扰决策方法在雷达任务较多的实战条件下决策效率会明显下降,因此文献[6]使用深度Q神经网络(Deep Q Network,DQN)来应对MFR任务较多的问题,但DQN算法存在计算Q值中“过估计”的问题^[7]。为此,本文采用Double DQN算法,通过不同的网络执行动作的评估与选择,对这两个部分进行解耦来解决这一问题,对空战中的干扰样式选择进行研究。仿真结果显示了本文算法的有效性。

1 干扰矩阵

现代空战已经进入了超视距时代,AFCR已经成为空战中主要的作战对象,发挥着不可替代的作用。AFCR要在不同的作战环境中遂行多样化的作战任务,因此在雷达系统设计上会设定多种工作模式,不同的工作模式对应着不同的威胁等级。空战中使用较多的是空-空模式下的雷达工作模式,因此本文以AFCR的空-空雷达工作模式为主开展干扰研究。

在典型的空战场景中,我方针对敌方的干扰实施流程如下:首先,我方通过侦察对敌AFCR的工作模式进行识别^[8],之后通过干扰样式选择算法从干扰样式集中选择最优干扰样式对其进行干扰,使其工作模式不断转换,直至转换为对应的威胁等级最低的状态,达到预期干扰效果。

根据雷达的工作模式及其转换关系和干扰样式集可以构建出干扰矩阵,将干扰矩阵作为算法输入

进行干扰样式选择。

1.1 典型空-空雷达工作模式

空-空模式下AFCR的工作模式^[9]主要包括搜索、跟踪、格斗和识别等,每一类模式下又有多种子模式。

1.1.1 搜索模式

速度搜索(Velocity Search,VS)模式主要用于远距高速接近目标的早期预警,威胁等级低。

边搜索边测距(Range While Search,RWS)模式可以在探测目标的同时进行测距,从而快速获取潜在目标的位置信息。

边扫描边跟踪(Track While Scan,TWS)模式^[10]主要用于对多个来袭目标的探测和跟踪。在检测到目标后火控计算机建立跟踪文件,再通过关联算法来对目标未来的参数进行估计。

1.1.2 跟踪模式

态势感知(Situation Awareness Mode,SAM)模式是RWS模式下的一个子模式,可以在单目标跟踪(Single Target Tracking,STT)的方式下同时保持对多个目标的探测,威胁等级较高。

高优先级目标(High Priority Target,HPT)模式是TWS模式的一个子模式,在TWS模式探测到的目标中指定一个为HPT,威胁等级较高。

单目标跟踪模式是AFCR具有最高威胁等级的模式。在选定HPT、SAM模式以及空战格斗模式下的子模式均可进入STT模式。

1.1.3 空战格斗模式

空战格斗(Air Combat Mode,ACM)模式主要用于敌我近距离空战,具有三种子模式。

机炮(Gun Acquire,GACQ)模式下,AFCR的平视显示器(Heads Up Display,HUD)上会出现一个20°视角的虚线圆圈。垂直(Vertical Acquire,VACQ)模式下,HUD上会出现两条距离5°视角的垂直虚线。这两种模式如果探测到一个位于5 n mile内的目标,将自动转入STT模式。

准星(Boresight,BST)模式下,HUD上的水线位置会出现一个3.4°视角的虚线圆圈,如果探测到目标并位于10 n mile内,将自动转入STT模式。

1.1.4 识别模式

非合作目标识别(Non-cooperative Target Recognition,NCTR)模式^[11]是载机通过AFCR利用非合作目标自身的反射特性对其进行远距离识别与分类。

1.2 干扰样式集

雷达干扰样式多种多样,每种干扰样式的原理又不尽相同,且对于不同的雷达工作模式作用效果也有所差别。我方机载干扰设备在作战前会进行干扰样式的预加载,用于在作战中的干扰样式选择。本文在干扰样式集的建立上选择了部分典型干扰样式^[12],包括相参和非相参干扰样式共 10 种:0 是密集假目标;1 是窄带噪声调相;2 是灵巧噪声;3 是相参梳状谱;4 是间歇采样;5 是距离-速度联合拖引;6 是多普勒噪声;7 是扫频噪声;8 是全脉冲复制;9 是移频干扰。这些干扰样式产生的干扰效果包括压制、欺骗和压制-欺骗联合效果。

1.3 干扰矩阵构建

干扰矩阵的构建可以分为实时建立和非实时建立。实时建立是在作战环境下,由机载侦察设备对

敌方 AFCR 的工作状态进行侦察与识别,根据我方做出干扰样式选择之后敌方 AFCR 的模式变化来获取雷达的工作模式转换关系,再结合我方的干扰样式集建立干扰矩阵。这种方式几乎没有任何先验知识,干扰样式是否可以起到效果无法得知。非实时建立则是在非作战环境下,依靠侦察设备在平时的战略侦察,我方已对敌方 AFCR 的工作模式有一定了解,通过对侦察数据的情报分析,对其在一些干扰样式作用下的工作模式转换关系已有一定的先验知识,再通过仿真手段对其他干扰样式进行验证、更新,得到一个较为完善的干扰矩阵。

本文的干扰矩阵是在第二种方式下建立的,并且结合了部分先验知识和开源资料。如表 1 所示,左侧一栏为威胁等级从高到低排列的工作模式,右侧则是各种干扰样式作用下的工作模式转换关系。

表 1 干扰矩阵表

| AFCR 工作模式 | 干扰样式 | | | | | | | | | |
|--------------|------|-----|------|-----|-----|-----|------|------|------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| STT | | HPT | | BST | | SAM | | STT | | |
| BST | | | GACQ | | | | VACQ | | | STT |
| GACQ | | BST | | | RWS | TWS | | GACQ | | RWS |
| VACQ | TWS | | RWS | | RWS | | BST | | VACQ | |
| NCTR | RWS | | | RWS | | | NCTR | | | |
| SAM | STT | | RWS | | | | RWS | | STT | |
| HPT | | TWS | | | HPT | | | TWS | STT | |
| TWS | | RWS | | | TWS | | | HPT | | RWS |
| RWS | RWS | | SAM | TWS | | | VS | | | SAM |
| VS | | | | | | | | | | |

2 基于 Double DQN 的干扰样式选择算法

2.1 DQN 原理与 Double DQN 的改进

2.1.1 DQN 原理

DQN 算法^[13]是一种从高维的输入学习控制策略的深度强化学习算法。和 Q-学习算法相比,DQN 将 Q 表的迭代更新过程转化为一个函数拟合问题,通过用函数来代替 Q 表产生 Q 值。

在普通的 Q-学习中,Q 值更新公式如下:

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma Q(s', a') - Q(s, a)). \quad (1)$$

式中: s 代表智能体 (Agent) 当前状态, a' 是 Agent 在下一个状态 s' 下选择的动作, r 表示 Agent 到达下一个状态得到的奖励值, $\alpha \in [0, 1]$ 是 Q 值更新中的学习率, $\gamma \in [0, 1]$ 是 Q 值更新中的折扣因子。 $Q(s, a)$ 为 Q 现实值, $r + \gamma \max_{a'} Q(s', a')$ 为 Q 估计值,通过对两者之间的误差学习来更新 Q 现实值。

在 DQN 算法中使用神经网络来对 Q 值进行估计,相近的状态得到相近的输出动作。通过更新网络参数 θ 来使 Q 函数逼近最优 Q 值,如式 (2) 所示:

$$Q(s, a; \theta_i) \approx Q(s, a). \quad (2)$$

式中: θ_i 是第 i 次迭代时的神经网络参数。

DQN 中有两个神经网络,两个网络的结构相同但参数不同。其中当前网络(MainNet)用来计算当前状态动作对的值函数 $Q(s, a; \theta_i)$,目标网络(TargetNet)的输出用来计算 Q 估计值中的 $Q(s, a; \theta_i^-)$ 。计算 Q 估计值的数学表达式如下:

$$\text{Target } Q = r + \gamma Q(s', a'; \theta_i^-) \quad (3)$$

式中: θ_i^- 表示 TargetNet 的参数。

损失函数如下:

$$\text{Loss}(\theta) = E[(r + \gamma Q(s', a'; \theta_i^-) - Q(s, a; \theta_i))^2] \quad (4)$$

MainNet 的参数是每次训练均在更新,每隔一定步数将网络的参数值赋给 TargetNet。在单步更新中 TargetNet 参数值不变, Q 估计值也不变。

2.1.2 Double DQN 的改进

DQN 中计算 Q 估计值时的最大化操作会使得估计的值函数比值函数的真实值大,即会产生非均匀的“过估计”,影响最终决策。作为一种离线学习的算法, DQN 每次学习时不是使用下一次交互的真实动作,而是使用当前认为价值最大的动作来更新目标值函数,但是对于真实的策略来说在给定的状态下并不是每次都选择使得 Q 值最大的动作,所以目标值直接选择动作最大的 Q 值往往会导致目标值要高于真实值。

在 Double DQN 算法中动作的选择和动作的评估分别用不同的值函数网络来实现:首先,动作的评估是根据 MainNet 得到下一状态 s' 下的 Q 值最大的动作 $Q(s', a'; \theta_i)$;然后,将这个动作作为计算 Target Q 值的动作,用 TargetNet 完成最优动作的选择工作。Double DQN 中 Target Q 的计算式如下:

$$\text{Target } Q = r + \gamma Q(s', Q(s', a'; \theta_i); \theta_i^-) \quad (5)$$

将上式代入损失函数计算公式,得

$$\text{Loss}(\theta) = E[(\text{Target } Q - Q(s, a; \theta_i))^2] \quad (6)$$

2.2 Double DQN 干扰样式选择方法概述

本文将 Double DQN 的原理应用于干扰样式选择,其算法模型的流程如图 1 所示,其中 $s \in S$ 表示敌方 AFCR 在某时刻的工作状态, S 代表敌方 AFCR 的工作状态集; $a \in A$ 表示干扰机在某时刻下选择出的干扰样式, A 代表我方干扰样式集。当我方侦察设备通过雷达工作状态识别得到 s ,将此状态输入到 Double DQN 网络中,经过神经网络拟合得到各个干扰样式对应的 Q 值,根据 ϵ -greedy 策略(该策略

可以平衡探索与利用之间的关系来获得累计最大回报)来选取出一种干扰样式 a ,作用于敌方 AFCR。根据敌方 AFCR 的工作状态转换,我方进行效果评估得到回报 $r \in \mathbb{R}$ 。由于敌方 AFCR 与我方属于非合作关系,我方可以根据遭受干扰后的雷达工作状态的改变来确定干扰的有效性^[14],通过对 AFCR 的威胁等级判定来给出相应的回报值 r 。干扰机根据回报值来更新 Q 值,决定下一工作状态 s' 所要采取的干扰样式,然后将上述的四个变量存入四元组样本 $\langle s, a, r, s' \rangle$ 中,多条样本组成样本池 D 。在样本池 D 中随机抽取一定批量(minibatch)的样本对神经网络进行训练,对网络参数进行更新,然后再进行干扰样式选择,循环往复,直至达到终止状态。

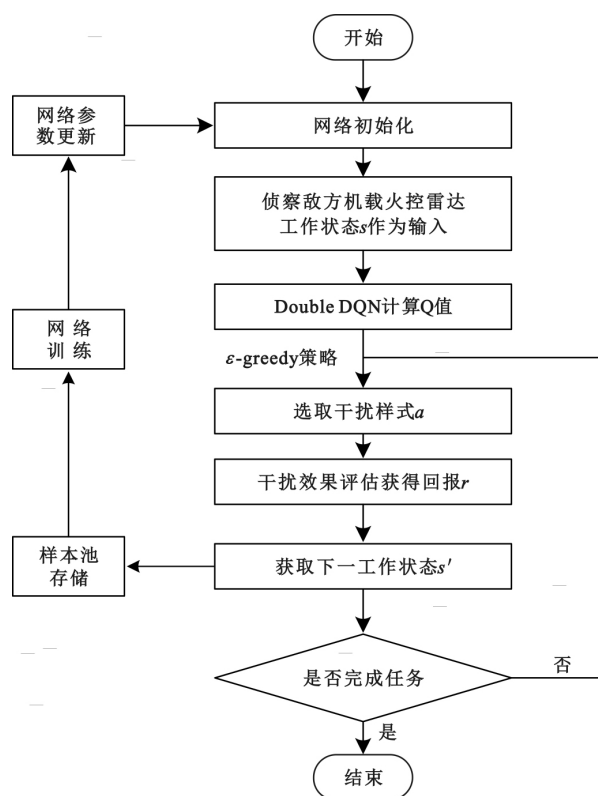


图 1 算法流程图

算法的伪代码如下:

```

1  输入: 状态空间  $S$ , 动作空间  $A$ , 折扣因子  $\gamma$ , 学习率  $\alpha$ , 迭代次数  $M$ 
   FOR  $i = 1: M$ 
2  初始化网络参数
   FOR  $s \in S$ 
3  输入侦察到的敌方 AFCR 工作状态  $s$ 
4  Double DQN 计算  $Q$  值并依据  $\epsilon$ -greedy 策略给出干扰样式  $a$ 

```

5 我方通过侦察得到 AFCR 下一工作状态 ,进行干扰评估得到奖励值 r

6 将四元组 存入样本池 D 随机抽取 minibatch 个样本对网络进行训练 ,更新网络参数

7 s 为终止状态;

END FOR

8 达到最大迭代次数 M

END FOR

由于不同的干扰样式产生的干扰效果不同 ,压制性干扰样式会使敌方 AFCR 的工作状态朝着威胁等级降低的方向转变;欺骗性干扰样式会产生欺骗性干扰效果 部分干扰样式可能会使其工作状态的威胁等级朝着升高的方向变化。因此 ,回报值定义如下:

$$r = \begin{cases} 100, & TL \rightarrow \min \\ +1, & TL \downarrow A_{Sj} \text{ or } TL \uparrow A_{Dj} \\ -1, & TL \uparrow \text{ or } TL \nleftrightarrow TL \end{cases} \quad (7)$$

式中: $TL \rightarrow \min$ 表示 AFCR 的工作状态威胁等级将至最低; $TL \downarrow A_{Sj}$ 表示采取压制性干扰样式 ,工作状态向威胁等级降低的方向转换; $TL \uparrow A_{Dj}$ 表示采取部分欺骗性干扰样式时工作状态威胁等级向升高方向转换; $TL \nleftrightarrow TL$ 表示工作模式之间没有转换; $TL \uparrow$ 表示 AFCR 工作状态向威胁等级升高的方向转换。

3 仿真验证

本文仿真使用 Tensorflow 框架编写。计算机配置如下:处理器采用 Intel(R) Core(TM) i5-10200H;内存为 16 GB。

本文用两个全连接层来近似 Q 函数 $Q(s, a; \theta_i)$ 。网络的细节如图 2 所示 ,该网络由两个全连接层组成。

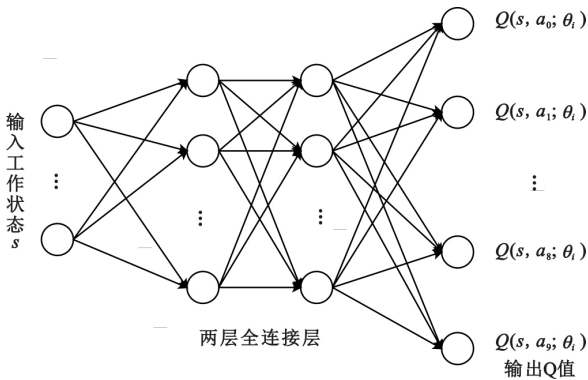


图 2 值函数拟合神经网络结构

算法的训练参数如表 2 所示。

表 2 参数设置

| 参数 | 取值 |
|----------------------------|-------------|
| 折现因子 | 0.9 |
| 学习率 | 0.005 |
| 样本池容量 | 1 000/2 000 |
| Batchsize | 64 |
| 网络参数交换频率 | 5 |
| 迭代次数 | 10 000 |
| ε -greedy 起止概率 | [0.1 0.9] |
| ε -greedy 增加幅度 | 0.001 |

在两个网络的参数替换前先观察 300 步 ,然后每隔 5 步将 MainNet 的参数赋给 TargetNet; 初始化的 ε -greedy 策略中 ε 的初始概率值为 0.1 终止概率值为 0.9 ,即初始时有 90% 的概率随机选择干扰样式 ,最终有 90% 的概率选择最优干扰样式 ,每一次学习都会将 ε 的值增大 0.001 ,增加到终止值时停止 这样的设置可以在训练的过程中使网络逐渐收敛; 每次从样本池的 2 000 个样本中随机抽取 64 个样本用于网络训练。

图 3 是 Double DQN 算法训练得到的损失图。由于样本量较少 ,在迭代中选择隔 5 步交换一次网络参数。在每隔 5 步交换网络参数后 ,网络训练损失值逐步下降 ,但由于初始时刻的 ε 探索值较低 ,随机选择干扰样式会导致损失值产生波动 ,但随着训练的进行 ε 探索值逐渐变大 ,神经网络拟合的 Q 值越来越准确 ,损失值在大约 750 步以后降至 0 附近 ,说明网络训练完成。整个训练过程所花费的平均时间为 11.7 s ,但在实际应用时 ,网络训练完成之后即可进行干扰决策 ,所以在实际应用中可以适当减少迭代次数 ,这样可以提高决策效率 ,使决策时间控制在秒级 ,满足实战要求。

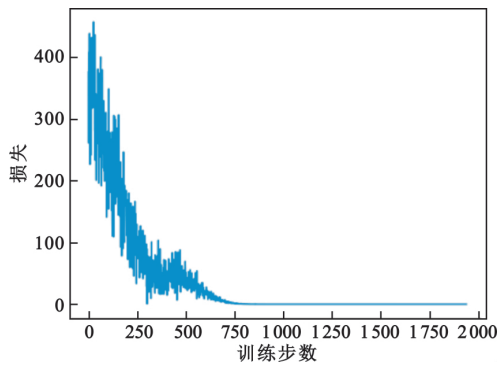
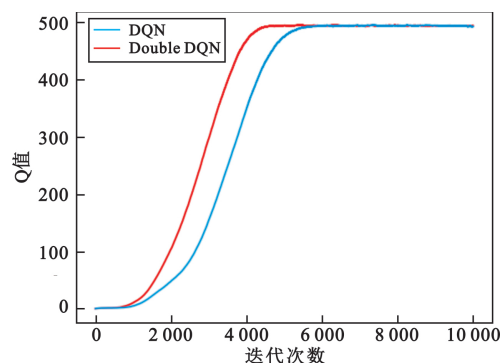


图 3 网络训练损失

将训练中每次用于干扰样式选择的 Q 值记录下来,如图 4 所示,可以看到两种算法的 Q 值变化趋势一致,在迭代 5 500 步后收敛至 495 附近。但是 Double DQN 算法的 Q 值消除了一部分过估计,相比 DQN 算法的 Q 值较低,计算 Q 值的平均值可以得到前者比后者低 34.01。这也验证了 Double DQN 算法在解决 Q 值偏大和网络训练的 Target Q 值不准确的问题上具有更良好的表现。

图 4 Q 值对比

强化学习算法的最终目标都是使总的目标奖励最大化,因此把迭代过程中每次干扰样式选择产生的奖励值相加得到总的奖励值。对比 20 次重复实验中 DQN 与 Double DQN 两种算法得到的总奖励值,如图 5 所示。可以发现,在绝大部分情况下 Double DQN 算法产生的总奖励值要高于 DQN 算法,这说明在干扰样式选择中选择最优动作比次优动作会获得更高的奖励值,从而验证了本文方法的有效性。

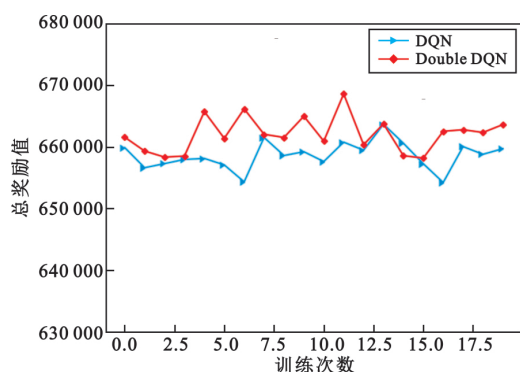


图 5 两种算法的总奖励值对比

在参数选择上,折扣因子 γ 是调节长期回报与即时回报的参数,在仿真中我们更注重长期回报,因此将 γ 设置为 0.9;而学习率 α 一般取值为 0.1,但由于本文仿真迭代次数较多,而且 α 值过大会引起损失函数剧烈波动,所以经过多次试验将 α 值设为 0.

• 1376 •

005; 样本池的容量没有固定的设置,当训练步数恒定为 2 000 步时分别设置样本池容量为 1 000 和 2 000,对比两种设置上的损失值情况,如图 6 所示,可见样本池容量设为 2 000 时损失值明显高于容量为 1 000 时的。因此,在训练步数为 2 000 步时选择样本池容量为 1 000。

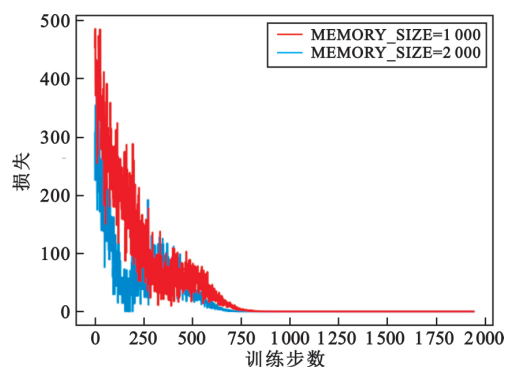


图 6 样本池容量对比

根据仿真结果可以得到如图 7 所示的最优干扰样式选择路径。图中以数字代表具体的干扰样式,根据空战的态势和 AFCR 雷达的工作模式分为近距和远距两类,分别如图 7(a)和图 7(b)所示。在不同的干扰样式的作用下,AFCR 的威胁等级由高到低逐渐下降;部分干扰样式在不同的态势和工作模式下产生的干扰效果也不尽相同。

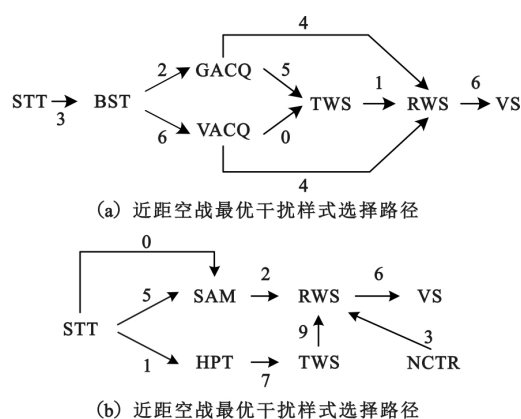


图 7 最优干扰样式选择路径图

4 结束语

随着机载火控雷达技术体制的不断进步,现代空战中敌我双方的电子战博弈也愈演愈烈。针对空战场景中机载火控雷达与干扰机之间的对抗,本文提出了一种基于 Double DQN 干扰样式选择方法,在建立干扰矩阵的基础上通过 Double DQN 算法选择最优干扰样式,仿真实验分析了参数设置,对比

DQN 证明了其在解决“过估计”问题上的有效性,给出了不同态势下的最优干扰样式选择路径。本文方法也可以为机载自卫吊舱的干扰样式选择提供参考。考虑到目前先进战机均为多用途战机,还可以执行对地、对海突击任务,因此下一步研究中还需对 AFCR 更多作战模式下的干扰样式选择问题进行探索。

参考文献:

- [1] 李丽.国外下一代战斗机及机载火控雷达发展[J].电子世界,2019(7):84-85.
- [2] 张柏开,朱卫纲.对多功能相控阵雷达干扰决策方法综述[J].兵器装备工程学报,2019,40(9):178-183.
- [3] 李云杰,朱云鹏,高梅国.基于 Q-学习算法的认知雷达对抗过程设计[J].北京理工大学学报,2015,35(11):1194-1199.
- [4] 邢强,贾鑫,朱卫纲.基于 Q-学习的智能雷达对抗[J].系统工程与电子技术,2018,40(5):76-80.
- [5] 张柏开,朱卫纲.基于 Q-Learning 的多功能雷达认知干扰决策方法[J].电讯技术,2020,60(2):129-136.
- [6] 张柏开,朱卫纲.对多功能雷达的 DQN 认知干扰决策方法[J].系统工程与电子技术,2020,42(4):819-825.
- [7] HASSELT H, GUEZ A, SILVER D. Deep reinforcement learning with double Q-learning [EB/OL]. [2021-03-01]. <https://arxiv.org/pdf/1509.06461.pdf>.
- [8] 马珂,毕大平,胡立群,等.基于 ELINT 的机载火控雷达空空工作状态识别[J].电子信息对抗技术,2019,34(4):35-39.
- [9] 唐玉文,何明浩,韩俊,等.机载火控雷达典型空-空工作模式浅析[J].现代防御技术,2018,46(6):87-93.
- [10] 欧健.机载火控雷达典型工作模式功率管理建模研究[D].长沙:国防科学技术大学,2013.
- [11] 朱自谦,刘睿.机载远距非合作式目标识别(NCTR)技术综述[J].空载雷达,2004(1):5-13.
- [12] 张柏开,朱卫纲.MFR 认知干扰决策体系构建及关键技术[J].系统工程与电子技术,2020,42(9):1969-1975.
- [13] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518(7540):529-533.
- [14] 赵耀东,徐旺.一种基于雷达状态变化的干扰效果在线评估方法[J].电子信息对抗技术,2016,31(3):42-46.

作者简介:



陈泽盛 男,1997 年生于山西怀仁,2019 年获工学学士学位,现为硕士研究生,主要研究方向为雷达干扰和强化学习等。

杨承志 男,1974 年生于甘肃甘谷,2013 年获工学博士学位,现为教授,主要研究方向为认知电子战和信息感知与对抗。

曹鹏宇 男,1997 年生于山西大同,2019 年获工学学士学位,现为硕士研究生,主要研究方向为认知侦察和深度学习。

邴雨晨 男,1989 年生于吉林长春,2014 年获工学硕士学位,现为讲师,主要研究方向为电子对抗与认知干扰。

纳 贤 男,1997 年生于云南昆明,2019 年获工学学士学位,现为情报分析师,主要研究方向为航空电子侦察。