基于 MDP 的 Massive MIMO 物理层安全算法

蒋 华[1,2] 侯梦茹[2] 张昕然[1] 王庆瑞[1]

[1](北京电子科技学院通信工程系 北京 100070)

[2](西安电子科技大学通信工程学院 陕西 西安 710071)

摘 要 对 Massive MIMO(multiple input multiple output)系统的物理层安全问题进行研究,提出基于马尔科夫决策过程 MDP(Markov decision process)的物理层安全模型。利用动态规划 DP(dynamic programming)求解 MDP 模型。利用互阻抗模型建立 Massive MIMO 系统下行链路的信道模型;建立基于平均无折扣回报的 MDP 模型,利用有限状态马氏信道 FSMC(finite state Markov channel)的区间转移概率模型,给出信道的转移概率表达式;提出基于值迭代算法的 MDP 物理层安全算法,得到加密容量最大化的全局最优策略;通过仿真对算法性能进行验证。仿真结果给出基站发射信号功率对信道物理层加密容量的影响,并评估算法的性能。

关键词 Massive MIMO 加密容量 MDP DP

中图分类号 TP393 **文献标识码** A DOI:10.3969/j.issn.1000-386x.2018.11.054

MASSIVE MIMO PHYSICAL LAYER SECURITY ALGORITHM BASED ON MDP

Jiang Hua[1,2] Hou Mengru[2] Zhang Xinran[1] Wang Qingrui[1]

[1](*Department of Communication Engineering, Beijing Electronic Science and Technology Institute, Beijing 100070, China*)

[2](*School of Telecommunications Engineering, Xidian University, Xi'an 710071, Shaanxi, China*)

Abstract We studied the physical layer security of massive MIMO(multiple input multiple output) system, and proposed the physical layer security model based on MDP(Markov decision process). The MDP model was solved by DP(dynamic programming). We established the channel downlink model of massive MIMO system by the mutual impedance model, and set up the MDP model based on average no-reward returns. The transfer probability expression of channel was given according to the interval transfer probability model of FSMC(finite state Markov channel). We also presented a MDP physical layer security algorithm based on value iteration to obtain the global optimal strategy with maximum encryption capacity. The performance of the algorithm was verified by simulation. The simulation results show the influence of the base station transmitting signal power on the channel physical layer encryption capacity, and evaluate the performance of the algorithm.

Keywords Massive MIMO Encryption capacity MDP DP

0 引 言

随着智能终端的普及,人们越来越依赖无线网络进行重要信息的传输。与通过密码技术保护数据安全的传统方法相比,通信系统的物理层安全通过利用通信介质的缺陷来提供安全的无线传输,不仅不依赖于计算复杂性,同时具有很高的可扩展性,为信息的保密传输提供了巨大优势[1-2]。作为最常用的物理层安全技术,MIMO 技术可以在一个或多个非法用户存在的情况下支持高速率的安全通信[3]。近年提出的大规模多天线阵列 MIMO 技术可以在不增加带宽或提高发射功率的情况下显著提高数据吞吐量和链路可靠性,因此成为 5G 移动通信系统的关键技术之一[4-5]。

在 Massive MIMO 系统中,在发射机上使用非常大的天线阵列(通常为数十甚至数百个)接收器,数百个

收稿日期:2018-04-13。**蒋华**,教授,主研领域:通信与信息安全。**侯梦茹**,硕士生。**张昕然**,讲师。**王庆瑞**,硕士生。

天线同时服务于数十个用户。理论和测量结果表明，大规模 MIMO 技术可以通过利用低复杂度传输设计提供的大阵列增益来提供高功率和能源效率。此外，当大量天线部署在基站时，可以降低随机损伤（如小规模衰落和噪声）的干扰[6-7]。由于 MIMO 技术只能辐射天线阵固定下倾角水平方向的波束，为了更好地利用信号传播的垂直角度分辨率，将 MIMO 的辐射信号控制在 3D 空间中，采用矩形、球形或圆柱形的天线阵列配置，被称为 3D MIMO 技术。第三代合作伙伴计划（3GPP）指出，具有大量天线的 3D MIMO 可被看作是 Massive MIMO 的实际形式之一[5]。

对于通信系统而言，信道容量上限代表了用户可达的最大速率。文献[8]提出了窃听信道三端口网络的加密容量的概念。窃听者通过其自己的通道收听传输信号不能解码消息的情况下，能够以严格正确的速率可靠地进行通信。即在加密容量存在的情况下，一定存在一种编码方式使得非法用户的信道容量为零。同时研究表明，随着天线阵列规模的增加，信道容量将和天线数呈线性增长的关系。因此对 MIMO 信道的物理层保密问题引起了研究学者的广泛兴趣[9]。

马尔可夫决策过程（MDP）模型是在不确定情况下进行顺序决策、考虑当前决策的结果和未来的决策机会的数学方法[10]。在近几年关于 Massive MIMO 技术的文献中，经常使用 MDP 模型作为优化工具，在多小区 Massive MIMO 系统中实现全局最优。例如使用 MDP 模型处理 Massive MIMO 通信系统的资源分配问题[11]。在文献[12]中，采用 SMDP（semi-Markov decision process）方法，提出了一种资源分配方案，以实现 OFDMA（orthogonal frequency division multiple Access）多小区协作网络中保证通信质量业务的最优功率效率。Massive MIMO 系统中的功率和速率分配问题在文献[13]中被建模为 CMDP（constrained Markov decision process），其优化目标是受延迟约束的最小化发射功率。下行链路 OFDMA 系统的功率和子载波分配问题在文献[14]中被建模为 CMDP，优化目标是在平均延迟约束下的最大化能量效率。当模型参数不可知时，强化学习算法被经常用来求解 MDP 模型。它是从控制理论、统计学、心理学等相关学科发展而来，具有自学习和在线学习的优点[15]。MDP 模型的最优策略可以用值迭代和动态规划算法确定。值迭代算法的优势在于其在实现上的简易性，可以用来进一步研究分析得到的最优策略的结构。

本文从物理层角度提出了基于 MDP 模型的 Massive MIMO 系统安全传输技术。利用互阻抗模型建立了 Massive MIMO 系统下行链路的信道模型。建立了基于平均无折扣回报的 MDP 模型，利用有限状态马氏信道 FSMC 的区间转移概率模型[16]，给出了 MDP 的转移概率表达式，并提出了基于值迭代的动态规划算法，计算了全局最优的系统加密容量。通过仿真对算法性能进行了验证，给出了仿真结果，并对结果进行了分析。仿真结果给出了基站发射信号功率对信道物理层加密容量的影响并评估了算法的性能。

1 系统模型

如图 1 所示，系统模型是多小区 Massive MIMO 系统的下行链路模型，此小区包括 K 个移动用户，每个小区的基站装载 N_T 根天线，每个用户有 N_R 根接收天线。考虑小区中使用相同时频资源的一个合法用户和一个窃听用户，则系统的物理场景简化模型如图 2 所示。

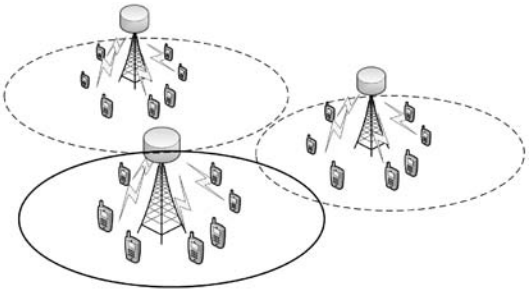

图 1 多小区 Massive MIMO 系统模型

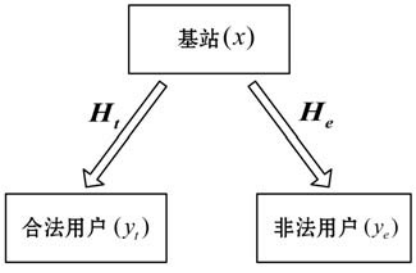

图 2 系统下行链路简化模型

系统模型为：

$$\boldsymbol{y}_{k,t} = \boldsymbol{H}_t \boldsymbol{x} + \boldsymbol{n}_t \tag{1}$$

$$\boldsymbol{y}_{k,e} = \boldsymbol{H}_e \boldsymbol{x} + \boldsymbol{n}_e \tag{2}$$

式中：$x \in \mathbb{C}^{N_T \times 1}$ 表示基站端的发射信号，$y_{k,t} \in \mathbb{C}^{N_R \times 1}$ 和 $y_{k,e} \in \mathbb{C}^{N_R \times 1}$ 分别表示第 k 个小区的合法用户和非法用户接收到的信号。$n_t \sim CN(0, \sigma_t^2)$ 和 $n_e \sim CN(0, \sigma_e^2)$ 为服从独立高斯分布，具有零均值和单位方差的加性高斯白噪声。$\boldsymbol{H} = \{\boldsymbol{h}_k\}$ 为信道传输矩阵。在 Massive MIMO信道场景下，应使用互阻抗模型对信道进行建模[5]。因此，第 k 个小区的传输信道为：

$$\boldsymbol{h}_k = \boldsymbol{g}_k \cdot \beta_k^{1/2} \tag{3}$$

式中：$\boldsymbol{h}_k \in \mathbb{C}^{N_R \times N_T}$。

$$\beta_k = \phi d_k^{-\alpha} \xi_k \tag{4}$$

$$\boldsymbol{g}_k = [\boldsymbol{Z}\boldsymbol{R}_k\boldsymbol{v}_k]^{\mathrm{T}} \tag{5}$$

式中:β_k表示大尺度衰落系数;ϕ 是与天线增益和载频有关的常数;d_k是第 k 个用户与基站之间的距离;α 是路径损耗指数;ξ_k是服从 $10\log_{10}\xi_k \sim N(0,\sigma_{sh}^2)$的阴影衰落分布。$\boldsymbol{g}_k$表示小尺度衰落系数矩阵,其元素服从独立高斯分布,具有零均值和单位方差,即 $g_k \sim CN(0,1)$;$\boldsymbol{v}_k$是一个随机向量,$\boldsymbol{v}_k \in \mathbb{C}^{N_R\times 1}$,且 $\boldsymbol{v}_k \sim CN(0,\boldsymbol{I}_{N_R})$,互阻抗矩阵 $\boldsymbol{Z} \in \mathbb{C}^{N_T\times N_T}$满足:

$$\boldsymbol{Z}\boldsymbol{P} = (A_Z + L_Z)(\boldsymbol{\Psi} + L_Z\boldsymbol{I})^{-1} \tag{6}$$

$$\boldsymbol{\Psi} = \begin{pmatrix} A_Z & M_Z & 0 & \cdots & 0 \\ M_Z & A_Z & M_Z & \cdots & 0 \\ 0 & M_Z & A_Z & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & M_Z & A_Z \end{pmatrix} \tag{7}$$

式中:A_Z、L_Z、M_Z分别表示天线的阻抗、负载阻抗和互耦阻抗。矩阵 $\boldsymbol{\gamma}_k \in \mathbb{C}^{N_T\times N_R}$,满足:

$$\boldsymbol{\gamma}_k = \frac{1}{N_T}[\boldsymbol{\zeta}(\theta_{k,1}),\boldsymbol{\zeta}(\theta_{k,2}),\cdots,\boldsymbol{\zeta}(\theta_{k,D_k})] \tag{8}$$

$$\boldsymbol{\zeta}(\theta_{k,i}) = [1,\mathrm{e}^{(j2\pi\,\tilde{d}/\lambda)\sin\theta_{k,i}},\cdots,\mathrm{e}^{(j2\pi(N-1)\,\tilde{d}/\lambda)\sin\theta_{k,i}}]^{\mathrm{T}} \tag{9}$$

式中:$\tilde{d}$ 为相邻天线之间的距离,λ 为载波波长。

2　基于 MDP 的值迭代算法

2.1　MDP 建模

为了模拟系统的物理层时变特性,建立有限状态马尔科夫信道(FSMC)模型来描述信道的时变行为。把下行链路增益量化为许多区间,第 j 个区间 ϕ_j对应一个链路增益范围:$\phi_j = \{\varphi:\chi_j \leqslant \varphi \leqslant \chi_{j+1}\}$,$\chi_j$为区间边界,信道被量化为 FSMC 模型。为上述 Massive MIMO 系统建立马尔科夫决策过程(MDP)模型,模型的组成部分包括 $<s,a,p,c>$,各项分别代表状态、动作、状态转移函数、回报函数,各项表述为:

(1) 状态 s:用 s_t表示 t 时刻下的信道状态,$s_t = \|h_k^H h_k\|$,它代表小区中用户的信道链路增益。马尔科夫决策过程在 t 时刻所有可行的信道状态构成一个状态集,称为状态空间 S_t,$S_t = \{s_1,s_2,\cdots,s_{N_T}\}$。

(2) 动作 a:动作用来控制系统的状态。a_t表示 t 时刻下 MDP 模型的动作,它代表基站的发射功率。马尔科夫决策过程在 t 时刻所有可行的行为构成一个行为集,即行为空间 A_t,$A_t = \{a_1,a_2,\cdots,a_{N_T}\}$。

(3) 状态转移函数 p:在离散的时刻 t,对状态 s_t采取动作 a_t,状态转移至下一状态 s_{t+1},其转移通过状态转移函数得到。文献[16]指出,假定在时间间隔 T 内 φ 值保持在同一个区间内,在这个时间间隔结束时,可能继续停留在本区间内或者转移到相邻的链路增益区间。定义状态增量函数 $\delta(\cdot)$用来表示相邻状态的变化量,则区间之间的转移概率为:

$$P(s'|s,a) = \begin{cases} p_{j,j+1} = L_{j+1}/\tilde{R} & s' = s+\delta(1) \\ p_{j,j-1} = L_j/\tilde{R} & s' = s-\delta(1) \\ p_{j,j} = 1 - p_{j,j-1} - p_{j,j+1} & s' = s \end{cases} \tag{10}$$

式中:L_j是χ_j处的电平通过率,即单位时间内信号包络向下穿过电平χ_j的平均次数,满足:

$$L_j = \sqrt{2\pi\chi_j/\rho}f_m\mathrm{e}^{-\chi_j/\rho} \tag{11}$$

式中:f_m是多普勒频率;ρ 为基站发送端信噪比的期望。

$$\rho = E\{SNR_t\} \tag{12}$$

$\tilde{R}$代表当前状态下单位时间内信号传输速率,满足:

$$\tilde{R} = R_t \times \kappa_j \tag{13}$$

式中:R_t代表单位时间内信号传输速率;κ_j代表当前状态下的稳态概率,满足:

$$\kappa_j = \int_{\chi_j}^{\chi_{j+1}} \frac{1}{\rho}\mathrm{e}^{-x/\rho}\mathrm{d}x = \mathrm{e}^{-\chi_j/\rho} - \mathrm{e}^{-\chi_{j+1}/\rho} \tag{14}$$

(4) 回报函数 c:在与环境的交互过程中,在离散的时隙 t,对状态 s_t采取动作 a_t,状态转移至下一状态 s_{t+1},产生回报。在此 Massive MIMO 系统中,以系统物理层的加密容量作为 MDP 模型的回报函数。由于 Massive MIMO 的物理层安全优势,可获得加密容量的表达式,从而无需使用任何正式的加密系统,该速率就可以可靠而安全地传输[17]。因此,三端口网络窃听系统中的加密容量为:

$$c(s,a) = \log\det\left(\boldsymbol{I} + \frac{1}{\boldsymbol{\sigma}_t^2}\hat{\boldsymbol{H}}_t\boldsymbol{R}_{xx}\hat{\boldsymbol{H}}_t^H\right) - \log\det\left(\boldsymbol{I} + \frac{1}{\boldsymbol{\sigma}_e^2}\hat{\boldsymbol{H}}_e\boldsymbol{R}_{xx}\hat{\boldsymbol{H}}_e^H\right) \tag{15}$$

式中:$\boldsymbol{R}_{xx} = E\{xx^H\}$,是发送信号的自相关矩阵。

(5) 策略 π:给定一个 MDP 模型,马尔科夫策略就是在某一状态下,决策者所采取的动作或者所采取的动作的概率。而在有限 MDP 中,一定至少存在一个策略 π,使得任意状态 $s_t \in S$ 下,值函数 $J_\pi(s_t) \geqslant J_{\pi'}(s_t)$,被称为最优策略 π^*。简单来说,解决一项强化学习任务的本质是寻找到最优策略。

2.2　基于值迭代的动态规划算法

强化学习算法是以评估价值函数为基础,通过价值函数将 MDP 的最优标准与策略联系起来。动态规划算法是在已知 MDP 模型的基础上,首先计算状态值函数,然后利用模型,计算出该状态下的最优动作,寻

找出最优化策略。为了计算系统最大化加密容量，采用值迭法算法，从初始状态价值开始反复迭代计算，最终收敛至全局最优价值函数 J^*，从而达到系统模型的最优结果。

以小区的加密容量作为整个系统的回报函数 $c(s,a)$，将问题转化为系统最优化问题，J_π^* 为马尔科夫决策过程的决策优化目标函数：

$$J_\pi^* = \lim_{T\to\infty}\sup\frac{1}{T}E\{\sum_{t=0}^{T-1}c(s_t,a_t)\} \tag{16}$$

根据贝尔曼方程，满足值迭代算法，最优值函数满足的迭代形式如下：

$$J_{l+1}(s,a) = c(s,a) + \max_{a'}\{\sum_{s'}P(s'\mid s,a)J_l(s',a')\} \tag{17}$$

式中：l 为迭代步数。对于每个状态 s，迭代地更新每一个状态动作对应的值，得到下一值函数 $J_{l+1}(s,a)$。直到 $J_{l+1}(s,a)$ 达到最优，满足：

$$J_\pi^* = \lim_{l\to\infty}J_{l+1}(s,a) \tag{18}$$

具体算法见算法 1。

算法 1 基于 MDP 模型的物理层安全算法

步骤 1 输入转移概率 $P(s'|s,a)$ 和奖赏函数 $R(s,a)$

步骤 2 初始化参数：令 $J_0(s,a)=0, temp=J_l(s,a)$

步骤 3 for 每个迭代步 $l=\{0,1,2,\cdots\}$

do

$J_{l+1}(s,a)=c(s,a)+\max_{a'}\{\sum_{s'}P(s'\mid s,a)J_l(s')\}$

$J_l(s,a)\leftarrow J_{l+1}(s,a)$

until

$|J_l(s,a)-temp|<\varepsilon$，$\varepsilon$ 为足够小量

步骤 4 $J_\pi^*\leftarrow J_l(s,a)$

3 仿真与分析

通过使用 MDP 工具箱对系统模型进行性能验证。由于 Massive MIMO 系统中的基站通常具有成百上千的天线，仿真难度太大，为了简化模型，对减少天线数的小规模 MIMO 系统进行了仿真，其中重点对策略求解进行了仿真。仿真参数设置为：多普勒频率 f_m = 5 kHz，信号传输速率 R_t = 200 Mbit/s，路径损耗因子 α 为 4.2，阴影衰落 σ_{sh}^2 = 10 dB。MDP 模型的状态参数，即信道增益的取值范围分别为：$s_1=\{30,40,50,60,70,80,90\}$，$s_2=\{50,60,70,80,90\}$，单位：dB。行为参数，即基站发射功率的取值范围为：$a=\{20,22,24,\cdots,86,88,90\}$，单位：dBm。

为了便于观察性能，定义信道的衰减值 $\Gamma = SNR_r^t/SNR_t$，其中 SNR_t 为基站发射端的信噪比，SNR_r^t 为合法用户接收端的信噪比，SNR_r^e 为非法用户接收端的信噪比，单位：dB。仿真结果如图 3 所示。

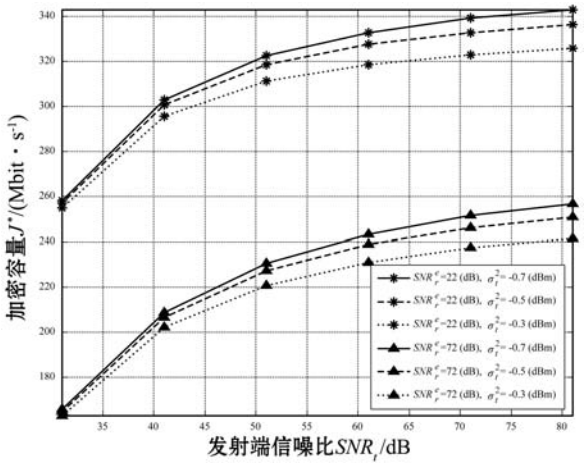

图 3 发射端信噪比对信道容量的影响

图 3 描述了基站发射信号对信道容量的影响，可以看出，随着发射端信噪比的增大，加密容量会增大。当最大发送信噪比大于某个门限信噪比时，策略的性能增长趋势减缓并趋于恒定，因此当发射功率较大时，即使再增加发射功率，也不能进一步提高系统加密容量。同时，窃听用户和信道噪声功率会影响加密容量，窃听用户接收信噪比增大，加密容量会减小；信道噪声功率增大，加密容量减小。因此通信环境需要警惕干扰信号，需要尽量减小干扰信号对加密容量的削弱作用，比如非法窃听用户的接收信号和信道噪声功率的负面影响，从而保证高质量的通信。

由图 4 可以看出，随着窃听非法用户接收端信噪比增大，加密容量会下降。非法用户小规模的窃听对加密容量的影响不大，当窃听用户窃听信号过大时，信道性能急剧恶化，因此通信传输过程中，减少非法用户的窃听是保证通信质量的基础。同时，图 4 体现出信道本身的衰减对加密容量也有影响。同等前提条件下，合法用户信道噪声功率的增加会削弱加密容量。

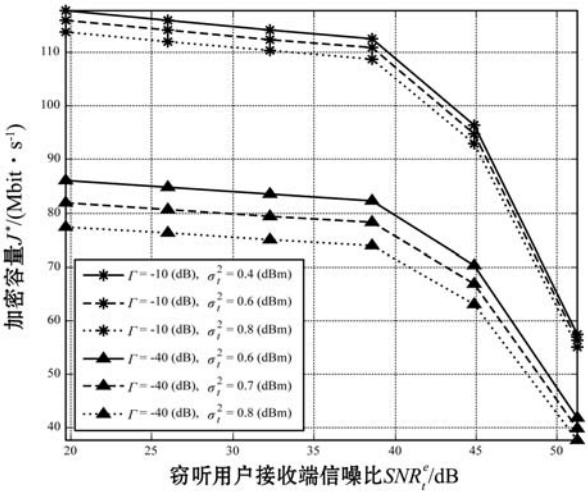

图 4 窃听用户接收端信噪比对信道容量的影响

表 1 是 MDP 模型的策略表,是一个输出动作标号的序列。它描述了算法在不同状态下寻找到的最优策略,即基站在每一状态下选择的行为以确保模型中的加密容量最大。具体来说,就是在此算法下,基站可以根据不同的信道增益状态选择发射功率,使系统模型获得最大的加密容量。

表 1　策略 π

策略 π 编号	行为 a 的标号
1	2
2	2
3	3
4	7
5	7
6	12
7	16
⋮	⋮
30	20
31	24
32	31
33	32
34	1
35	1

4　结　语

本文基于 MDP 模型研究了 Massive MIMO 系统的物理层安全算法。通过对 Massive MIMO 下行链路模型进行 MDP 建模,利用值迭代算法求解 MDP 模型,计算出系统最大化加密容量。根据算法模型,基站可以在不同信道增益下控制发射功率,求解出保证物理层加密容量的全局最优策略。分析结果说明了基站发射信号功率对三端口网络信道物理层加密容量的影响,即系统的加密容量会随着基站发射功率的增大而增大,随着非法窃听用户的接收信噪比的增大而减小,证明了算法的正确性和有效性。进一步说明强化学习的相关算法可以解决通信系统中相关的控制问题,为今后的研究奠定了理论和应用基础。

参　考　文　献

[1] Leng S, Ng D W K, Schober R. Power efficient and secure multiuser communication systems with wireless information and power transfer[C]//IEEE International Conference on Communications Workshops. IEEE, 2014:800 - 806.

[2] Liu L, Zhang R, Chua K C. Secrecy wireless information and power transfer with MISO beamforming[J]. IEEE Transactions on Signal Processing, 2014, 62(7): 1850 - 1863.

[3] Malkowsky S, Vieira J, Liu L, et al. The world's first real-time testbed for massive MIMO: design, implementation, and validation[J]. IEEE Access, 2017, 5: 9073 - 9088.

[4] Yang N, Wang L, Geraci G, et al. Safeguarding 5G wireless communication networks using physical layer security[J]. Communications Magazine IEEE, 2015, 53(4):20 - 27.

[5] Zheng K, Ou S, Yin X. Massive MIMO channel models: A survey[J]. International Journal of Antennas & Propagation, 2014, 2014(Article ID 848071):1 - 10.

[6] Larsson E G, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless system[J]. IEEE Communications Magazine, 2014, 52(2):186 - 195.

[7] Harris P, Beach M, Armour S, et al. From MIMO to massive MIMO[J]. Microwave Journal, 2017, 60(9):22 - 42.

[8] Wyner A D. The wire-tap channel[J]. Bell Labs Technical Journal, 1975, 54(8):1355 - 1387.

[9] Bjornson E, Hoydis J, Sanguinetti L. Massive MIMO has unlimited capacity[J]. IEEE Transactions on Wireless Communications, 2017,17(1): 574 - 590.

[10] Weng P, Spanjaard O. Functional reward markov decision processes: theory and applications[J]. International Journal on Artificial Intelligence Tools, 2017, 26(3):1760014.

[11] Li P, Jiang Y, Li W, et al. A CMDP-based approach for energy efficient power allocation in massive MIMO systems [C]//Wireless Communications and NETWORKING Conference. IEEE, 2016.

[12] Wang P, Zhang X, Song M. Optimal stochastic subcarrier and power allocations for QoS-guaranteed services in OFDMA multicell cooperation networks[C]//IEEE International Conference on Communications. IEEE, 2013:6449 - 6453.

[13] Djonin D V, Krishnamurthy V. MIMO transmission control in fading channels—a constrained markov decision process formulation with monotone randomized policies[J]. IEEE Transactions on Signal Processing, 2007, 55 (10): 5069 - 5083.

[14] Bi K, Yang Q, Fu F L, et al. Energy-efficient power and subcarrier allocation for OFDMA systems with value function approximation approach [C]//International Conference on ICT Convergence. IEEE, 2012:530 - 535.

[15] 刘全. 大规模强化学习[M]. 北京:科学出版社, 2016.

[16] Goldsmith A. Wireless Communications[M]. 北京:人民邮电出版社, 2007.

[17] Kapetanovic D, Zheng G, Rusek F. Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks[J]. IEEE Communications Magazine, 2015, 53(6):21 - 27.