

# 基于强化学习的无线可充电传感网移动充电路径优化



张昊 管昕洁 白光伟

南京工业大学计算机科学与技术学院 南京 211816

(Becky951219@126.com)

**摘要** 无线传感器网络在环境感知、目标跟踪等方面占据了重要地位。为了能够及时地为传感器节点补充能量,提出了一种基于强化学习的低功耗、高能效的移动路径充电算法。无线传感器网络采用移动充电车对传感器节点进行充电,将 Q-Learning 算法与 epsilon-greedy 算法相结合,以最短路径依次完成所有传感器节点的充电。现有的相关研究通常忽略了传感器节点自身所能承受电量的最大值,容易导致传感器节点因充电过程中电量超出最大值而暂停工作,因此限制了移动充电车的充电时间。结果表明,所提移动充电策略的效用更高,与传统的 Q-Learning 算法和贪心算法相比,训练周期大幅度下降且实现了能量利用率最大化。

**关键词**: 无线可充电传感网;移动充电;路径;强化学习;能量利用率

**中图分类号** TP393

## Optimization of Mobile Charging Path of Wireless Rechargeable Sensor Networks Based on Reinforcement Learning

ZHANG Hao, GUAN Xin-jie and BAI Guang-wei

Department of Computer Science and Technology, Nanjing University of Technology, Nanjing 211816, China

**Abstract** Wireless sensor networks occupy an important position in environmental perception and target tracking. In order to recharge sensor nodes in time, this paper proposes a low power consumption and high energy efficiency mobile path charging algorithm based on reinforcement learning. Wireless sensor network uses a mobile charger to charge the sensor nodes. The Q-Learning algorithm and the epsilon-greedy algorithm are combined to complete the charging of all sensor nodes in turn in the shortest path. Existing related researches usually ignore the maximum amount of power that the sensor node itself can withstand, which easily causes the power to exceed the maximum threshold during charging and suspend work, so the charging time of the mobile charger is limited. The result shows that the proposed mobile charging strategy has a higher utility. Compared with the traditional Q-Learning algorithm and the greedy algorithm, the training cycle is greatly reduced and the energy utilization rate is maximized.

**Keywords** Wireless rechargeable sensor network, Mobile charging, Path, Reinforcement learning, Energy utilization

### 1 引言

随着无线技术的快速发展,无线传感器网络在现实生活中占据了愈来愈重要的地位,尤其在检测和监测应用中发挥了重要作用,包括环境感知、目标跟踪和结构健康监测等<sup>[1]</sup>。传统的无线传感器网络(Wireless Sensor Network, WSN)中,传感器节点一般由电池驱动,而电池有限的能量不仅阻碍了传感器的大规模部署,还影响了整体网络的工作寿命。针对这一问题,研究人员提出了能量节省<sup>[2]</sup>和能量汲取<sup>[3]</sup>的方法,试图减少或平衡传感器的能量消耗,但均无法从本质上解决这一问题。就能量节省方法而言,它只能实现网络寿命的有效延长,并不能阻止传感器节点耗尽能量,网络最终还是死亡;而能量汲取方法虽然通过传感器节点能够从周边环境

汲取大量能量,但是考虑到环境的不可测性,汲取效率会深受影响,导致传感器节点不能持续、稳定地工作。

Kurs 等<sup>[4-5]</sup>提出可采用无线能量传输技术(Wireless Power Transfer, WPT)来解决传统 WSN 中传感器节点电池能量受限的问题,使得无线可充电传感网(Wireless Rechargeable Sensor Network, WRSN)得到了进一步发展。WRSN 的首要问题是无线充电器的布置。目前,已有工作对此问题展开了研究。Dai 等<sup>[6]</sup>在给定的区域内任意布置若干个有向充电器,并设置充电器的朝向和位置,实现了传感器网络的整体充电效用最大化。这是首次通过实测数据建立的有向充电模型,也是基于该模型首次考虑了有向无线充电器平面的一般布置问题。

然而,这种通过定点部署充电器的充电方式增加了研究

收稿日期:2020-04-17 返修日期:2020-07-13 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金项目(61802176)

This work was supported by the National Natural Science Foundation of China (61802176).

通信作者:管昕洁(xjguan@njtech.edu.cn)

成本,因而大量研究工作都利用移动充电器(Mobile Charger, MC)来为 WRSN 中的传感器节点补给电量,有效减少了能量源的部署开销,减少了能量耗散。MC 的移动速度、充电功率以及总电能是有限的,如何合理利用 MC 来维持 WRSN 长期稳定地工作逐渐成为近来研究的热点问题之一。

Shi 等<sup>[7]</sup>提出了可再生能源周期的概念,在每一个周期内尽可能将充电器的移动时间缩至最短,并将此类问题形式化为哈密尔顿回路问题。Xie 等<sup>[8]</sup>优化了充电器的移动策略,使得整体网络的能量损耗最小化。He 等<sup>[9]</sup>提出了一种充电策略,充电器按照最近任务优先的原则选择传感器进行充电,从而最小化传感器待充电的等待时间。Jiang 等<sup>[10]</sup>首次综合考虑了传感器的调度和充电器的移动,实现了网络检测性能的最大化。

然而,这些移动策略中 MC 通常需要遍历 WRSN 所有传感器节点进行充电,在传感器节点接收功率、能耗速率各异的实际情况下,遍历所有传感器节点会导致 MC 对高剩余能量传感器节点进行过度充电,使得这类传感器节点因负荷过大而暂停工作,这是因为其缺乏对传感器节点剩余能量阈值、存储电量峰值的考虑。本文重点讨论了如何规划 MC 的移动路径,控制充电时间,在提高 WRSN 整体充电效率的同时降低能量损耗。本文提出了一种高效低耗的移动充电策略,构建了传感器节点单元化的充电模型,并提出了充电路径的最小化问题,以实现更高效的移动充电,保证 WRSN 长期稳定地工作。

## 2 系统模型

### 2.1 网络模型

将 WRSN 看成二维平面<sup>[11]</sup>,随机布置 1 个基站和  $n$  个传感器节点,基站和传感器节点的位置均固定不变。为了方便表述,在二维平面上建立坐标轴,即可用坐标  $S=(x, y)$  表示传感器节点的具体位置,其中  $S=(s_0, s_1, s_2, \dots, s_n)$  是  $n$  个传感器节点的集合,则任意两个传感器节点间的距离表示为  $dis_{ij}$ ,其中  $i, j=1, 2, \dots, n$ 。

### 2.2 移动充电模型

传感器节点的初始信息  $Inf_{s_i}$  表示为  $Inf_{s_i}=(x_i, y_i, E_r)$ ,包含传感器节点的位置以及剩余电量  $E_r$ 。所有传感器节点的最大储能均相同,记为  $E_{max}$ 。

图 1 为一个 MC 为无线可充电传感网进行充电的一般模型示例图。MC 初始时位于基站,并以固定的速度移动,依次对传感器节点进行充电,随后返回基站并补给自身能量。

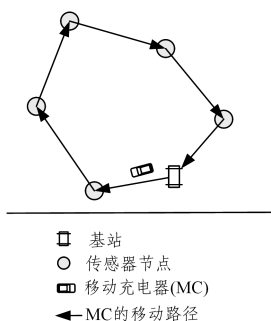


图 1 WRSN 一般充电模型的示例图

Fig. 1 Example of general charging model in WRSN

如图 2 所示,MC 的充电范围是一个半径为  $R_{ad}$  的圆形区域,即当 MC 对某一传感器节点进行充电时,以该传感器节点为中心、 $R_{ad}$  为半径的圆形区域内的所有传感器节点都将处于被充电状态。本文利用 WPT 能量传输的全向性,建立了单元化<sup>[12]</sup>充电模型,将传感器节点所在平面分割为面积相同的圆形区域,MC 只停留在单元中心处进行充电,进一步提高了移动充电的效率。

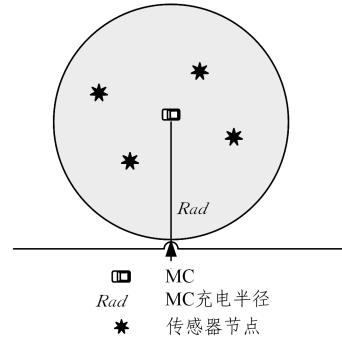


图 2 MC 充电范围示意图

Fig. 2 Schematic diagram of MC's charging range

就任意单元而言,传感器节点与中心位置的距离远近直接影响了接收功率的大小,其关系式<sup>[13]</sup>如下:

$$P_S(d) = \begin{cases} \frac{P_{MC}}{(d+\beta)^2}, & E_r < E_{max} \\ 0, & E_r = E_{max} \end{cases} \quad (1)$$

其中,  $P_S(d)$  表示与中心距离为  $d$  的传感器节点的接收功率,  $\beta$  为距离改变时的补偿参数,  $P_{MC}$  为 MC 的充电功率。当传感器节点的自身电量已经为最大值  $E_r = E_{max}$  时,则不需要进行充电,且接收功率为 0。考虑到中心传感器节点的充电情况,当  $d=0$  时,令  $\beta=1$ ,以确保关系式的有效性。

通常,单个传感器节点需要充的能量为传感器节点的最大能量与剩余能量的差值,记为  $en_i$ ,即:

$$en_i = E_{max} - E_r \quad (2)$$

因此可得:

$$en(c) = \sum_{i=0}^c en_i \quad (3)$$

其中,  $en(c)$  为充电过程中所有传感器节点所充能量的总和。

在单元化充电模式下,传感器节点的剩余电量各异容易使得部分传感器节点所接收的能量超出其自身的最大储能,即引发过充问题,从而停止工作,甚至影响整体网络的寿命。因此,须确保 MC 对传感器节点的充电量不超出传感器节点自身可承受的最大值。从上述接收功率的表达式可以看出,充电时间与每个传感器节点的剩余电量以及传感器节点与区域中心的距离有关,剩余电量最大并不直接表明充电时间最短。区域内每个传感器节点电量充满的时间可表示为:

$$t_i = \frac{E_{max} - E_r}{P_S(d)} \quad (4)$$

为避免出现过充问题,须满足每个区域的充电时间  $t$  不超过最小充电时间,即  $t \leq t_{min}$ ,使得所有传感器节点能够稳定地工作,传感器节点各自的剩余电量及时更新为:

$$E_{r_{new}} = tP_S(d) + E_r \quad (5)$$

小车自身的能量为  $E_{MC}$ ,供自身移动和为传感器节点充电。MC 的移动速度为  $V_{MC}$  且保持匀速,移动过程中的损耗

功率为  $P_{\text{损耗}}$ 。由于 MC 能量有限,因此在充电过程中,MC 给传感器节点所充的总能量和小车行驶所需的能量之和不能超过小车自身携带的电池能量<sup>[14]</sup>,约束公式为:

$$en(c) + E_{\text{损耗}} \leq E_{\text{MC}} \quad (6)$$

其中,  $E_{\text{损耗}}$  为 MC 在行驶中消耗的总能量:

$$E_{\text{损耗}} = \int E_{\text{损耗}}(t) dt = \int t P_{\text{损耗}} dt \quad (7)$$

因此,在小车能量有限的情况下,给予传感器节点最大化充电以保证网络的正常工作,同时选择最佳移动路径,使得 MC 的能量利用率达到最高。能量利用率具体表现为 MC 对传感器节点充电的总能量与移动过程中消耗的总能量之比,即:

$$\Delta = \frac{en(c)}{E_{\text{损耗}} + en(c)} \quad (8)$$

在满足以上约束条件的情况下,本文基于 Q-Learning 算法求解出 MC 的最优路径,由式(8)将该问题形式化为:

$$\begin{cases} \max \Delta = \frac{en(c)}{E_{\text{损耗}} + en(c)} \\ en(c) + E_{\text{损耗}} \leq E_{\text{MC}} \\ t \leq t_{\min} \end{cases} \quad (9)$$

实现能量利用率的最大化,需要满足两个条件:充电效用最佳和 MC 移动损耗能量最少。在不考虑外在环境的影响且 MC 匀速移动的情况下,MC 移动路径最短即可实现移动充电中 MC 的耗能最少。虽然 MC 移动充电路径的规划问题在形式上非常简单,但实际上却具有一定的挑战性。传感器节点分布的随机性为构建单元化充电模型带来了无限可能,传感器节点位置的不同间接改变了其接收功率和充电时间,从而影响了 MC 的移动方向和能量利用率。

### 3 移动充电算法

移动充电算法的目的是规划 MC 的最优路径,实现整体网络充电效用的最大化。该算法分为两部分:一部分是将传感器节点所在平面分割为若干个圆形单元,保证每个子区域能覆盖最多的传感器节点;另一部分则是规划 MC 的移动路径,遍历整体网络进行充电,实现 MC 的能量利用率最大化。

#### 3.1 单元的分割

平面上 MC 的充电位置有无限多个,给后续的有效处理带来了难度。本节利用类似于平面区域离散化<sup>[15]</sup>的思想,将整个平面划分为有限个圆形单元,充电位置为圆形单元的中心处,从而使得无限多个候选充电位置降为有限多个。如上所述,MC 的充电范围是一个半径为  $Rad$  的圆形区域,因此将平面划分为若干个半径为  $Rad$  的圆形区域。

以每一个传感器节点为中心、以  $Rad$  为半径,找出能同时进行充电的传感器节点,这可能会引起一个传感器节点重复出现在多个充电区域内的情况,因此做如下处理:统计初步划分的圆形充电单元中包含的传感器节点个数及传感器节点的信息,以最少的圆形区域包含最多的传感器节点为准则,删除相对无效的圆形区域,降低传感器节点的重复出现率,如图 3 所示。同时这种一对多的充电模式实现了更高效的移动充电。最后,获取  $K$  个有效单元及其中心坐标,即 MC 在各个单元的充电停留位置记为  $U_j, j=1, 2, \dots, K$ 。除此之外,将

基站作为一个独立单元。

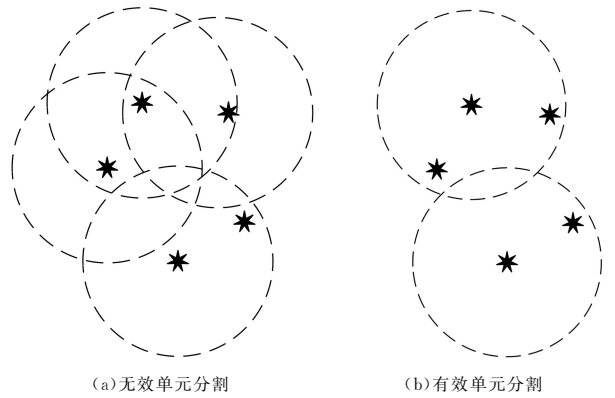


图 3 单元分割示意图

Fig. 3 Schematic diagram of unit division

#### 3.2 MC 的充电路径规划

MC 对任意单元充电后,自动选取下一个充电目标,依次对所有单元完成充电,即 MC 执行充电任务时需要进行自动决策,并且可以实现连续决策,找出最佳移动路径。因此,本文基于 Q-Learning 算法的基本思想来规划 MC 的最优充电路径,实现能量利用率最大化。

##### 3.2.1 Q-Learning 算法的概述

Q-Learning 的基本思想是:通过在环境中不断地进行尝试,根据尝试获得的反馈信息调整策略,多次训练后最终生成一个最佳策略,机器根据这个策略便可知当前状态下可选取的最优动作<sup>[16]</sup>。Q-Learning 的关键在于 Q-table, Q-table 通常以矩阵形式出现,行和列分别表示状态和采取相应行为的反馈值, Q-table 的值  $Q(s, a)$  用来衡量在当前状态下采取此行为是否是最佳选择。Q-table 中的数值更新公式如下:

$$Q(s, a) = r + \gamma(\max_{a'}(Q(s', a'))) \quad (10)$$

其中,  $s$  表示当前状态;  $a$  表示当前动作;  $r$  表示当前的收益;  $\gamma$  表示折扣因子,即时间的远近对回报的影响程度,一般取值在 0 至 1 之间;  $s'$  表示下一个状态;  $a'$  表示下一个动作;  $Q(s', a')$  则要在 Q-table 中查询<sup>[17]</sup>。

##### 3.2.2 Q-Learning 算法的改进策略

Q-Learning 算法的传统步骤为:首先从初始状态开始,并将其设置为当前状态;然后从当前状态出发,找到具有最高 Q 值的动作;接着执行该动作并找到下一状态,再将其设置成当前状态;重复以上步骤直到当前状态是目标状态为止。Q-Learning 算法中奖励值的设定也极为重要,传统的赋值方式如式(11)所示:

$$R = \begin{matrix} \text{action} & 0 & 1 & \dots & k \\ \begin{bmatrix} -1 & 0 & \dots & 100 \\ -1 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 100 & -1 & \dots & -1 \end{bmatrix} & \text{state}_0 \\ & & & & \text{state}_1 \\ & & & & \vdots \\ & & & & \text{state}_k \end{matrix} \quad (11)$$

其中,  $-1, 0, 100$  均表示当前状态下采取相应行动的奖励值。本文目的是求解出 MC 的最短移动路径,及时反馈。Q 值越大反映出当前移动距离越短。行动  $action_k$  表示从当前状态移动到下一状态  $state_k$ , 状态  $state_k$  表示 MC 停留在第  $k$  个单

元进行充电。合理选取一个常量  $\alpha$ , 根据欧氏距离公式<sup>[18]</sup> 计算当前状态下所在单元中心位置与其他单元中心位置的距离, 并将常量  $\alpha$  与距离的差值作为相应的奖励值, 即当前状态下采取行动至下一对应状态的及时反馈  $r$ , 直接表现为任意单元间距离的远近程度。

传统的 Q-Learning 算法求解最短路径使得 Q 值表接近收敛状态, 需要经过很多次的反复训练, 尤其是在传感器节点数量繁多且分布复杂的情况下, 因而降低了求解效率。为了缩短训练周期, 提高算法有效性, 本文在 Q-Learning 算法的基础上结合了 epsilon-greedy 方法来选取最优动作并加速算法的收敛, 具体流程如图 4 所示。

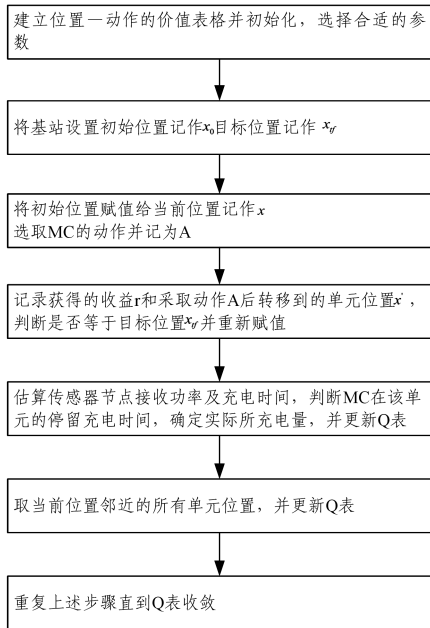


图 4 路径规划算法流程

Fig. 4 Flow chart of path planning algorithm

## 4 实验结果

为验证所述移动充电策略的有效性, MC 遍历 WRSN 对所有传感器节点完成一次充电, 将所述算法的训练周期数以及 MC 的能量利用率作为评价指标。

仿真环境设置如下: 在  $60 \times 60$  的二维平面随机布置  $n$  个传感器节点和 1 个基站, 位置固定不变。建立坐标轴并用坐标表示传感器节点的具体位置:  $S_i = (x_i, y_i)$ ,  $i = 1, 2, \dots, n$ 。计算任意两个传感器节点之间的距离  $dis_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ ,  $i, j = 1, 2, \dots, n$ , 将距离值存储于矩阵中。实验中的参数设置如表 1 所列。

表 1 实验参数设置

Table 1 Parameters setting in experiment

参数	值	描述
$n$	20, 30, 40, 50, ..., 100	传感器节点数量
$Rad$	3	MC 充电范围半径
$E_{max}$	1	传感器节点最大能量
$P_{MC}$	0.5	MC 充电功率
$\gamma$	[0, 1]	折扣因子

图 5 显示了坐标轴上传感器节点个数为 30、基站个数为

1 的随机分布情况。根据平面上传感器节点的分布情况, 以传感器节点间的距离为基准建立单元化充电模型<sup>[12]</sup>。如图 6 所示, 平面上共形成 7 个半径均为 3 的圆形单元, 将 30 个传感器节点全部覆盖, 且有两个传感器节点重复出现在两个不同单元内。

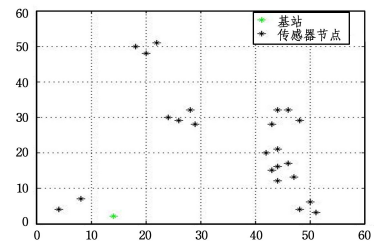


图 5 30 个传感器节点随机分布示意图

Fig. 5 Schematic diagram of random distribution of 30 sensors' nodes

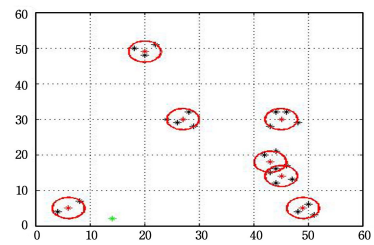


图 6 单元化分割示意图

Fig. 6 Schematic diagram of unitized segmentation

WRSN 中有一辆移动充电车从基站出发, 随后依次前往各个单元对传感器节点进行充电, 最终回到基站补给自身能量, 此为一轮充电调度。依据所述策略 MC 对每次移动的反馈值进行预估, 如图 7 所示, 随着训练周期数的递增, MC 的移动距离总体呈下降趋势, 直至 Q 值表接近收敛状态, 即得出当前分布情况下的最佳移动路径。

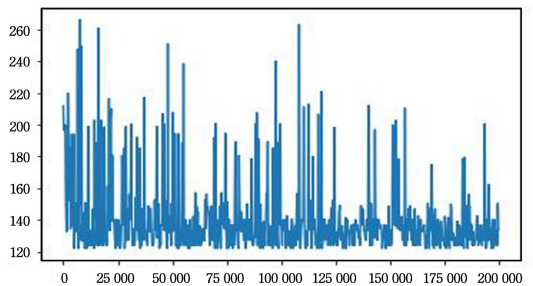


图 7 算法训练次数与 MC 路径长度的变化趋势

Fig. 7 Trend of algorithm training times changing with MC path length

如图 8 所示, MC 从基站 6 号圆形单元出发, 遍历 WRSN 对传感器进行充电, 最后返回到基站, 该路径为在此分布情况下的最佳移动充电路径。为了避免出现过充问题, 在对任意单元进行充电时, 需要计算该单元内所有传感器节点的接收功率, 并得出每个传感器节点理论上的充电时间, 从而控制 MC 在该单元的停留时间不超过传感器节点中的最短充电时间。

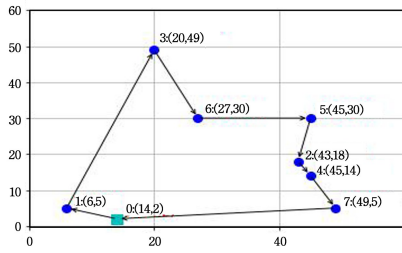


图 8 MC 充电移动路径示意图

Fig. 8 Schematic diagram of MC's moving path

结合图 6 可以看出,在单元 2 和单元 4 中存在两个相同的传感器节点。根据上述内容,在对任意单元充电完成后需要实时更新每个传感器节点的剩余电量  $E_{r_{\text{新}}} = tP_{\text{接收}}(d) + E_r$ ,因此在单元 2 充电完成后可获取重复出现传感器节点的最新剩余电量并进行判断,为单元 4 的充电任务作好充分的准备。随着传感器节点数量的增加及其随机分布的复杂性,传感器节点会重复存在于两个甚至两个以上的单元内,因此,及时更新传感器节点的自身电量有助于提高 MC 对充电时间把控的准确性并削弱单元间的互相干扰。

在较短的工作周期下,MC 以最短路径对每个单元都进行了一次充电且未出现路径重复的现象,降低了 MC 在移动上的能量开销,维持 WRSN 稳定工作的同时实现了总体能量利用率最大化。由能量利用率公式  $\Delta = \frac{en(c)}{E_{\text{损耗}} + en(c)}$  可以看出,传感器节点所获得的总能量与 MC 移动中损耗的能量是影响能量利用率的关键因素。移动路径最短实现了 MC 能量损耗最小,而传感器节点所获得的总能量是每个节点补充电量的总和。每个单元的充电时间  $t$  不能超过该单元内传感器节点中的最短充电时间,且各个单元内传感器节点的位置、数量各异,从而影响了各个单元整体所补充的电量及充电效率。

图 9 仿真了在传感器节点数为 20~100 时 MC 能量利用率的变化曲线。

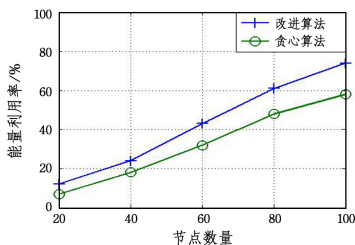


图 9 传感器节点数量对能量利用率的影响

Fig. 9 Impact of number of sensor nodes on energy utilization

贪心算法的思想是整体最优解可以通过一系列局部最优的选择来达到,即作出在当前看来是最好的选择但并不从整体最优的角度考虑。由图中的对比可以看出,在传感器节点数量较多且分布复杂时,根据贪心算法规划的充电路径会出现 MC 走回头路的现象,增加了 MC 在移动中的能量损耗,进一步降低了能量利用率。本文所述算法随着传感器节点数量的增加,在实现了移动路径总长度最短的前提下,MC 能够为更多的传感器节点提供充电服务。因此,能量利用率也在不断提高。

为了让求解的最优值更具有效性,在传感器节点数为 20~100 时进行多次仿真实验,取多组优化后训练周期数的平均值,图 10 为本文方法与传统 Q-Learning 算法的对比结果。由图 10 可知,在传感器节点数量和分布情况相同的条件下,采用 epsilon-greedy 选取最优行动与传统的 Q-Learning 算法相比,随着传感器节点数量的增多,求解路径需要的训练周期数呈现明显的减少趋势,Q 值收敛速度得到了显著提升。

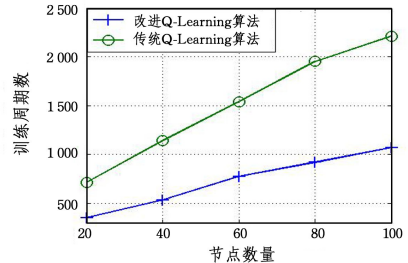


图 10 节点数量对训练周期数的影响

Fig. 10 Impact of number of sensor nodes on training cycles

综上所述,本文提出的改进算法降低了 MC 在移动中的能量损耗,实现了能量利用率最大化,同时减少了解决问题所需的训练周期,提高了算法的有效性。此外,该改进算法还考虑了所有传感器节点的充电时间以及自身电量,限制了 MC 在各个单元的最大充电时间,从而维持了传感器节点的正常运行,延长了网络整体生存周期,进一步提高了移动充电的整体能效。

**结束语** 本文研究了 WRSN 中的移动充电问题,构建了传感器节点单元化的充电模型,并针对该模型提出了一种高效低耗的移动充电策略来求解出最短充电路径。在求得最短路径的基础上,考虑了传感器节点的接收功率和能耗速率各异的实际情况,估算单元内所有传感器节点自身的充电时间以控制 MC 在各个单元的充电时间,避免出现传感器节点因负荷过大而暂停工作的问题。在传感器节点数量多且分布复杂的情况下,为保证求解效率,将 Q-Learning 算法与 epsilon-greedy 算法相结合,加速了算法的收敛,从而提高了算法的有效性。在一轮充电调度内,规划出的最短路径降低了 MC 在移动上的能量损耗,在维持 WRSN 稳定工作的同时实现了总体能量利用率的最大化。

## 参考文献

- [1] LÓPEZ RIQUELME J A, SOTO F, SUARDÍAZ J, et al. Wireless Sensor Networks for precision horticulture in Southern Spain[J]. Computers & Electronics in Agriculture, 2009, 68(1): 25-35.
- [2] GIUSEPPE A, MARCO C, MARIO D F, et al. Energy conservation in wireless sensor networks: A survey[J]. Ad Hoc Networks, 2009, 7(3): 537-568.
- [3] FAFOUTIS X, VUCKOVIC D, DI M A, DRAGON N, et al. Energy-Harvesting wireless sensor networks[C]//Proc. of the 9th European Conf. on Wireless Sensor Networks(EWSN). Trento: University of Trento, 2012: 84-85.
- [4] KURS A, KARALIS A, MOFFATT R, et al. Wireless power

- transfer via strongly coupled magnetic resonances[J]. *Science*, 2007, 317(5834):83-86.
- [5] KURS A, MOFFATT R, SOLJACIC M. Simultaneous midrange power transfer to multiple devices[J]. *Applied Physics Letters*, 2010, 96(4):34.
- [6] DAI H P, CHEN G H, XU L J, et al. Effective Algorithm for Placement of Directional Wireless Chargers[J]. *Ruan Jian Xue Bao*, 2015, 26(7):1711-1729.
- [7] SHI Y, XIE L, HOU Y T, et al. On Renewable Sensor Networks with Wireless Energy Transfer[C]// *INFOCOM*, 2011 Proceedings IEEE. IEEE, 2012:1350-1358.
- [8] XIE L, SHI Y, HOU Y T, et al. Multi-node wireless energy charging in sensor networks[J]. *IEEE/ACM Transactions on Networking*, 2015, 23(2):437-450.
- [9] HE L, KONG L, GU Y, et al. Evaluating the On-Demand Mobile Charging in Wireless Sensor Networks[J]. *IEEE Transactions on Mobile Computing*, 2015, 14(9):1861-1875.
- [10] JIANG F C, HE S B, CHENG P, et al. On optimal scheduling in wireless rechargeable sensor networks for stochastic event capture[C]// *IEEE International Conference on Mobile Adhoc & Sensor Systems*. IEEE, 2011.
- [11] SU Z Z. Research and Improvement on Routing Protocols of Wireless Sensor Networks Based on Clustering[D]. Changchun: Jilin University, 2016.
- [12] XIE L, SHI Y, HOU Y T, et al. Multi-Node Wireless Energy Charging in Sensor Networks[J]. *IEEE/ACM Transactions on Networking*, 2015, 23(2):437-450.
- [13] HE S B, CHEN J M, JIANG F C, et al. Energy provisioning in wireless rechargeable sensor networks[J]. *IEEE Transactions on Mobile Computing*, 2013, 12(10):1931-1942.
- [14] YU L C, LV H F, HE L, et al. Optimization of Charging Path for Wireless Rechargeable Sensor Networks [J]. *Journal of Shanghai DianJi University*, 2018(4):25-30.
- [15] FU L K, CHENG P, GU Y, et al. Minimizing charging delay in wireless rechargeable sensor networks[C]// *2013 Proceedings IEEE INFOCOM*. IEEE, 2013:2922-2930.
- [16] VARTIAINEN E M, INO Y, SHIMANO R, et al. Numerical phase correction method for terahertz time-domain reflection spectroscopy[J]. *Journal of Applied Physics*, 2004, 96(8):4171-4176.
- [17] GU S, LILICRAP T, SUTSKEVER I, et al. Continuous deep q-learning with model-based acceleration [C] // *International Conference on Machine Learning*. 2016:2829-2838.
- [18] SONG Y C, ZHANG Y Y, MENG H D. Research Based on Euclid Distance with Weights of Clustering Method[J]. *Computer Engineering and Applications*, 2007, 43(4):179-180, 226.



**ZHANG Hao**, born in 1995, postgraduate. Her main research interests include reinforcement learning, artificial intelligence and wireless sensor network.



**GUAN Xin-jie**, born in 1984, Ph.D, master instructor. Her main research interests include network optimization, edge computing and software defined network.