

基于 Q-Learning 的无线区块链广播路由算法^①

魏京生^② 司鹏搏^③ 李 萌 张延华

(北京工业大学信息学部 北京 100124)

摘 要 面向无线区块链网络场景,参考常用的区块链逻辑分层,提出一个独立灵活的分层逻辑架构,详细地介绍了每层的工作原理及作用。在现有广播路由算法的基础上结合无线区块链网络场景,针对广播过程中存在的广播冗余及广播总能耗高等缺陷,提出了一种基于 Q-Learning 的广播路由算法。该算法将无线区块链网络交易信息广播过程描述为多个 agent 的广播决策问题,通过线下学习、训练,帮助每一个当前广播节点 agent 做出是否广播的决策。仿真实验结果证明,该算法有效地解决了广播冗余问题,减少了广播过程的总能耗。

关键词 无线区块链; Q-Learning; 广播冗余; 能耗

0 引 言

区块链技术具有去中心化、公开透明、不可篡改等特点,其本质上是解决信任问题、降低信任成本的技术方案。近年来,区块链技术在数字货币、物联网等众多领域都得到了广泛的应用,极大程度地解决了安全、维护等问题^[1]。区块链中的交易可理解发生在对等点之间,而不是通过一个中央服务器或是中间处理器。每一笔交易信息会被广播到其他节点,且需全网承认有效,这一过程可实现区块链无法篡改、安全系数高、奖励诚实惩罚作恶的特点,解决信任问题,降低信任成本^[2],由此可知,广播机制在区块链网络中尤为重要。由众多节点组成的无线区块链网络^[3]进行交易信息的广播操作时,时延、可达率、能耗等无线网络中的通信性能指标也是无线区块链网络中的广播路由策略的服务质量(Quality of Service, QoS)指标,广播冗余、广播风暴等也是无线区块链网络中亟待解决的问题。此外,区块链全网广播时还需考虑节点故障、恶意节点作恶甚至是

非法入侵、伪造消息等原因导致的节点遗漏问题。

路由算法是提升网络性能的重要方式,因此必须设计一些有创新性且具备高效节能性的路由算法。目前无线网络大多采用动态路由算法以适应网络拓扑结构的变化及流量的变化。不同于单播的路由选择算法,广播路由算法需提供一种从源节点到网络中的所有其他节点交付分组的服务。由于简单且容易实现,传统的洪泛广播路由算法是无线网络中经常采用的一种广播算法,其基本思想是接收到信息的节点以广播的方式转发数据包,但在传统的洪泛实现过程中,存在着严重的广播冗余以及能源浪费问题。近年来,众多学者从降低无谓冗余及尽量避免广播风暴^[4]角度着手,提出各种算法。例如 Durresi 等人^[5]提出了一种分布式的节能广播算法——传感器网络广播算法(broadcast protocol for sensor networks, BPS),以降低转播冗余为目的,该算法扩大一次转播覆盖范围,从而一次转播可覆盖更多的节点。罗瑛等人^[6]对以节约能量为目的的广播路由算法进行研究,寻找节点能耗的平衡点,提出改进后的最长边距最小化(advanced minimum lon-

① 国家自然科学基金(61901011)资助项目。

② 男,1995年生,硕士生;研究方向:区块链、机器学习;E-mail: weijingsheng@emails.bjut.edu.cn。

③ 通信作者,E-mail: sipengbo@bjut.edu.cn。

(收稿日期:2021-03-05)

gest edge, AMLE) 广播路由算法, 该算法构造、维护一个具有最长边距最小化 (minimum longest edge, MLE) 性质的广播树, 在平衡各节点能耗的同时维持较低的广播过程总能耗。无线区块链网络大多采用 Gossip 协议实现洪泛广播^[1], 由此出现的问题如前所述, 该类广播算法引出的资源消耗巨大等问题一直在制约着区块链的发展。而目前众多学者在区块链广播领域内的研究大多集中于安全层面, 现有广播路由算法多数是从安全角度出发进行设计。例如, Sun 等人^[7]提出在区块链支持的无线物联网网络中最优全节点部署的算法对无线区块链系统的交易进行安全性能分析。Biswas 和 Dasgupta^[8]提出一种改进式的目的节点序列距离矢量 (destination-sequenced distance-vector, DSDV) 路由协议, 引入区块链技术来提供安全保护。然而区块链技术本身在内存、计算、功率上消耗巨大, 如若不考虑资源消耗的问题, 引入区块链技术的优势将会大打折扣。

本文首先参考常用的区块链逻辑分层, 针对无线区块链网络提出一个新的分层逻辑架构。此外, 对该分层逻辑架构中的网络层进行研究, 针对目前区块链网络经常采用的洪泛广播路由算法存在的缺陷, 以降低无线区块链网络广播冗余及广播过程总能耗为目的, 结合无线区块链网络交易信息广播时存在的问题, 引入强化学习领域的技术, 提出一种基于 Q-Learning 的无线区块链广播路由算法, 将无线区块链网络交易信息广播过程描述为多个 agent 的广播决策问题。本文将事先以放弃无谓广播降低冗余为目的设定的奖励值表 (Reward-Table) 及需做出是否广播决策的节点周边邻居信息等作为依据, 对网络中的节点进行多次线下学习训练, 最终得到一个收敛的 Q 值表 (Q-Table); 对于无线区块链网络中每一个需做出广播决策的节点, 判断其节点所处状态, 结合 Q-Table, 使其做出获得最大利益的决策。仿真结果表明, 所提出的算法相较于洪泛广播路由算法在能耗、可达率等指标上都有极大的改善。

1 无线区块链网络逻辑架构

参考传统的开放式系统互联 (open system inter-

connection, OSI) 模型和现有的区块链分层逻辑架构^[9], 本文设计了 6 层的系统分层逻辑架构来构建无线区块链网络。为保证灵活性, 每一层各司其职, 相互独立但又不可分割。例如作为基础的数据层, 其中包含的数据、交易记录、节点信息等是整个无线区块链体系的基础; 而数据层中又含有汇聚子层, 其提供与网络层对接的接口, 为网络层将节点与数据联通提供通道, 方便网络层实现其信息交互的相应工作。该架构如图 1 所示, 其主要内容如下。

(1) 应用层: 支持分布式应用 (decentralized application, DAPP)、资源配置、数据展示等应用。一方面可以使用下层数据, 另一方面对下层进行配置。

(2) 合约层: 合约层封装区块链系统的各类脚本代码、算法以及由此生成的更为复杂的智能合约, 当合约层中的智能合约达到其约束条件时自动触发执行。

(3) 安全层: 包含权限管理及数字签名, 以防止节点信息被恶意伪造及篡改。

(4) 共识层: 让高度分散的节点在去中心化的系统中高效地针对区块数据的有效性达成共识。区块链中比较常用的共识机制主要有工作量证明、权益证明和股份证明 3 种。



图1 无线区块链网络分层逻辑架构

(5)网络层:网络层的主要目的是实现区块链网络节点之间的信息交互,规定了整个网络节点里的通信机制,可以实现没有中心服务器的数据共享,包含以下 2 个子层。

P2P 子层:区块链网络的本质是一个 P2P(点对点)网络,包含节点的布局,保证不同类型的节点各司其职,形成高效的协同工作。

路由子层:一个节点接收到用户交易时,会广播到网络中并尽量传播到各个节点,一个节点生产出来的区块,也要求尽快传播到各个节点,此子层包含节点进行通信时采用的广播路由算法。

(6)数据层:包含汇聚子层和感知执行子层。

汇聚子层:将感知执行子层的数据结果进行汇聚整合,用于数据层和网络层的接口。

感知执行子层:包含了感知器、执行器、用户终端等数据的源和宿,进行信息感知与决策执行。

2 区块链广播中节点遗漏问题

拜占庭将军问题是区块链网络中的关键问题之一。当 $n > 3f$ 时(即叛徒的个数 f 小于将军总数 n 的 $1/3$ 时)将军们可以达成一致的命令,即在该范围内存在任意多的叛徒(至少得有 2 个忠诚将军),都可以找到解决方案。综上,无线区块链网络进行全网广播时,若广播过程中因节点故障、恶意节点作恶、网络故障等原因导致广播节点遗漏,可统称这些节点为“故障节点”,个数为 f (如图 2 所示)。如果全网总节点个数 $n \geq 3f + 1$,则可找到相应的解决方案^[10-11]。此外,当区块链网络中存在恶意节点数量超过网络中全部节点数量的 $1/3$ 但小于全网节点数量的半数(即 $(n-1)/3 < f < n/2$)时,仍存在一些达成一致的解决方案^[12]。但恶意节点数量超出全网节点的半数时,区块链系统本身会处于极不稳定的状态,此时对达成一致性带来极大困难。

假设广播遗漏的节点个数为 x ,如若将其一并视为故障节点,即包含在故障节点 f 中,当全网总节点个数 $n \geq 3f + 1$ 时,如前文所述,可找到达成一致的方法,由此可为广播过程中产生的个别节点遗漏情况提供解决方案。

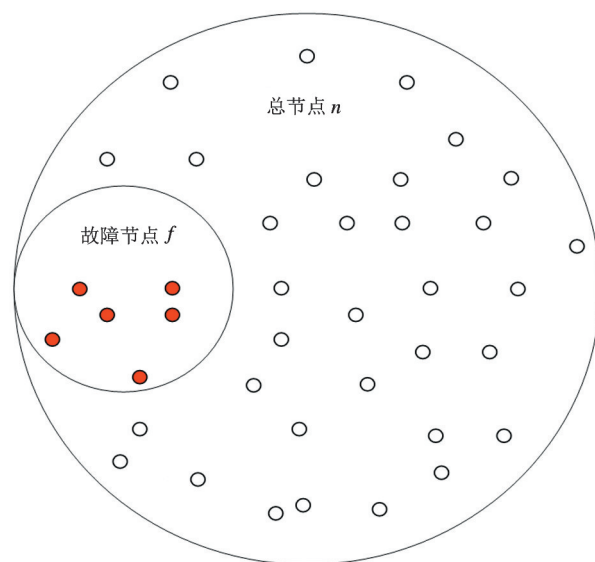


图 2 无线区块链网络中节点构成

综上所述,在研究无线区块链广播路由算法时,考虑到无线区块链网络可容纳一定数量的节点遗漏,同时也为更好地观测无线区块链广播路由算法性能,可将广播总体目标在不影响区块链系统状态及广播总体要求的前提下适当降低。

3 基于 Q-Learning 的无线区块链广播路由算法

针对上述问题,本文在分析无线区块链场景的特点及广播需求后,提出基于 Q-Learning 的广播路由算法,与传统的洪泛广播路由算法相比,该算法有效地解决了广播冗余问题并降低了广播过程总能耗。

在本文所提出的算法中,无线区块链网络中每一个节点都被视为一个 agent,每个 agent 通过判断周边邻居节点状态,做出是否进行广播的决策。通过不断的学习和取样, Q 值将会收敛,无线区块链网络广播可被描述为多个 agent 的强化学习问题^[13]。假设节点的发射功率保持稳定。

3.1 邻居节点信息维护

每个节点维护自己的邻居节点信息表如表 1 所示,该信息表由当前节点传输区域内每个邻居节点周期性地广播状态报文产生,其中包含节点传输区域内邻居节点总数 M ,其传输区域内已经接收到广

播信息的节点个数 m 及两者的比例;状态报文如表 2 所示,其中包含该发送节点的位置及节点状态(即是否已接收到广播信息(1/0))。当节点接收到状态报文时,可根据其中的节点位置做出判断。如若发送该状态报文的节点不在邻居节点信息表中,则将节点添加到邻居节点信息表中;如若发送该状态报文的节点存在于邻居节点信息表中,则可根据状态报文中包含的节点状态更新该发送节点的状态信息,更新邻居节点信息表中的信息。

表 1 邻居节点信息表的格式

节点 ID	邻居节点总个数	已经接到广播信息的节点个数	比例 P
ID	M	m	m/M

表 2 状态报文格式

节点 ID	节点状态	节点位置
ID	I/O	(location 1, location 2)

3.2 系统模型

对应本文提出的算法,每一个需做出广播决策的节点所处状态 s_i 可定义为无线区块链网络中节点的自身信息和相邻节点的信息,即当前需要进行做出是否广播的节点与周边所有邻居节点中已经接收到广播信息的节点比例 $P^{[14]}$ 。

动作 a_i 可定义为当前节点是否进行广播,即:

$$A = [a_1, a_2], a_i \in A \quad (1)$$

其中 a_1 表示“广播”, a_2 表示“不广播”。

综上,对于当前需要做出是否进行广播决策的节点 n_i ,通过查询其邻居节点信息表判断其状态;选择动作 a_i ,根据 Q 值评估状态-动作,选择最大 Q 值所对应的动作。

另外,节点进行一次广播的能量消耗的计算公式^[15]为

$$C_e(i) = lr^b + \frac{5}{3}C_e \quad (2)$$

其中 l 为节点发送数据的能耗系数, r 为节点传输半径, b 为环境影响系数, C_e 为节点处理信息的能耗,通常为常量。

广播过程总能耗^[16]为

$$C = \sum_T C_e(i) \quad (3)$$

其中 T 为在广播过程中做出“广播”决策的节点数量(转播节点个数)。

3.3 广播动作优化

Q-Learning 是一种记录行为值(Q-Value)的方法^[17],在一定的状态 s_i 下,每种行动 a_i 都会有一个值 $Q(s_i, a_i)$,其计算公式如下:

$$Q(s_i, a_i) = Q(s_i, a_i) + \alpha[R + \gamma \max_{a_{i+1}} Q(s_{i+1}, a_{i+1}) - Q(s_i, a_i)] \quad (4)$$

其中 $\gamma(0 \leq \gamma \leq 1)$ 是折扣因子,当 γ 接近 0 时,agent 更重视短期利益;而 γ 接近 1 时,agent 更重视长远利益, α 为学习率。Q-Learning 的核心思想是建立一个 Q 值表,记录在相应的状态下采取不同的动作得到的 Q 值,agent 可以根据 Q 值表来做出决策^[18]。

本文提出的基于 Q-Learning 的广播路由算法的原理是:事先根据拟划分的 S 种状态范围及 2 种可选取的动作建立奖励值表,奖励值表设定的目的为指引节点在广播时放弃无谓广播及降低冗余,因此在建立奖励值表时需针对不同节点状态,赋予不同动作不同奖励值。例如若节点状态 s_i 为 90% ~ 100%,则此时节点做出“广播”动作必然会产生广播冗余,造成不必要的能源浪费,因此此时“广播”动作对应的奖励值需为负值,“不广播”动作对应的奖励值为正值;同时需创建好一张 Q 值表,与奖励值表同阶,初始化为 0(Q 值代表当前节点选择是否进行广播所获得的期望回报,从而可以基于 Q 值表做出决策),之后经过多次线下训练,根据 Q 值计算公式更新 Q 值表,最终使其收敛。在无线区块链网络中,对于当前需要做出广播决策的节点,通过查询其邻居节点信息表判断其所处状态 s_i ,从而根据 Q 值表做出是否广播的决策,具体步骤如下:

(1) 初始化 Q 值表, $Q \leftarrow 0$

(2) 对应奖励值表,根据不同状态任意选取动作,训练学习,根据式(4)更新 Q 值表经过不断的采样和学习, Q 值表最终收敛。

(3) 对于所有节点,当需要其做出是否广播的决策时,通过查询其邻居节点信息表判断其状态 s_i ,

结合 Q 值表,做出能够获得最大收益的决策。

4 仿真与分析

本文使用 Matlab 模拟一个长宽均为 10 km 的矩形区域,且假设 100 个节点在这个区域中,由此构成整个无线区块链网络,并选取其中任意一个节点作为源节点。为防止在广播过程中由于遗漏节点对观测算法性能造成影响,同时也为适应上述分析中对无线区块链网络广播时的特点的分析,在仿真时以

$$S = \begin{bmatrix} 0 \sim 10\% , 10\% \sim 20\% , 20\% \sim 30\% , 30\% \sim 40\% , 40\% \sim 50\% \\ 50\% \sim 60\% , 60\% \sim 70\% , 70\% \sim 80\% , 80\% \sim 90\% , 90\% \sim 100\% , \\ 100\% \end{bmatrix} \quad (5)$$

本文设定奖励值表如表 3 所示。

表 3 奖励值表

动作 at	a_1	a_2
状态 S		
0 ~ 10%	20	-30
10% ~ 20%	15	-25
20% ~ 30%	13	-23
30% ~ 40%	11	-20
40% ~ 50%	9	-18
50% ~ 60%	7	-10
60% ~ 70%	6	-9
70% ~ 80%	4	-2
80% ~ 90%	0.5	-0.5
90% ~ 100%	-0.5	0.5
100%	-30	20

在理想状况下,广播过程总体成本主要集中于节点进行广播操作的总体能量消耗,节点周期性维护所涉及的传输成本和计算成本可忽略不计。此外,奖励值表中状态个数及相应动作奖励值是自主设定。如若状态个数划分较多,每种状态范围区间较为精细,则仿真结果在相应指标上的体现会更为精确,相应的学习训练、计算成本也会提高。

图 4 为节点传输半径对基于 Q-Learning 的广播路由算法及洪泛广播路由算法在广播过程中转播率(做出广播决策的节点个数与广播总体节点个数的

全网 90% 及以上的节点都接收到广播消息为全局目标,如图 3 所示。此外,每个节点的传输半径可从 1.5 km 至 2.0 km 自定义。传统的洪泛路由广播算法是一种常用的算法,因此本文在广播总能耗这一指标上将基于 Q-Learning 的广播路由算法与传统的洪泛广播路由算法做比较。本文设定环境影响系数为 4,能耗系数为 1。为方便计算,设定 C^e 为 0。此外,本文设定奖励值表状态范围共 11 种,以集合形式列出:

比例)的影响。由图 4 可知,随着节点传输半径增大,基于 Q-Learning 的广播路由算法与传统的洪泛路由算法在转播率这一指标上,前者一直低于后者。

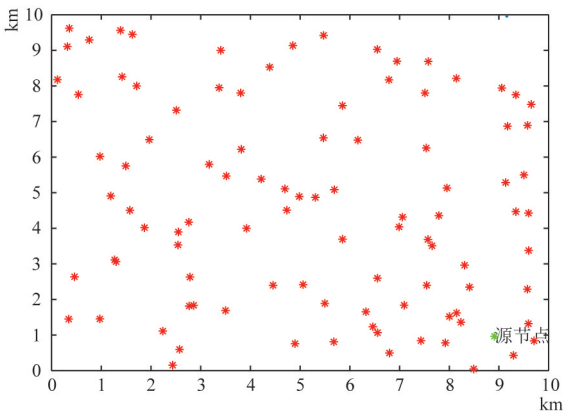


图 3 全网广播示意图

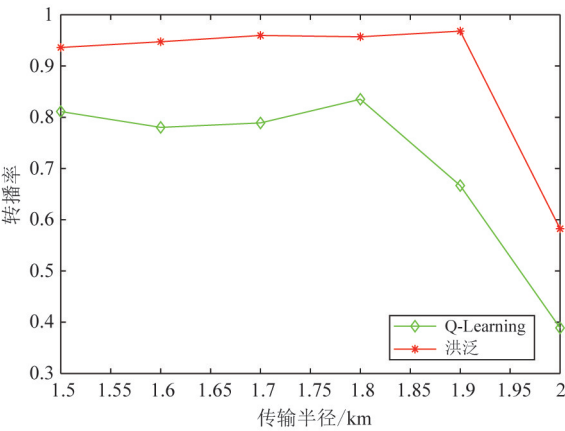


图 4 两种算法关于转播率的比较

图 5 为在达到全局目标(全网 90% 及以上的节点都接收到广播消息)的前提下,节点传输半径对基于 Q-Learning 的广播路由算法及洪泛广播路由算法在广播总体节点个数上的影响。由图 5 可知,随着节点传输半径增大,基于 Q-Learning 的广播路由算法与传统的洪泛路由算法在广播总体节点个数上基本持平。

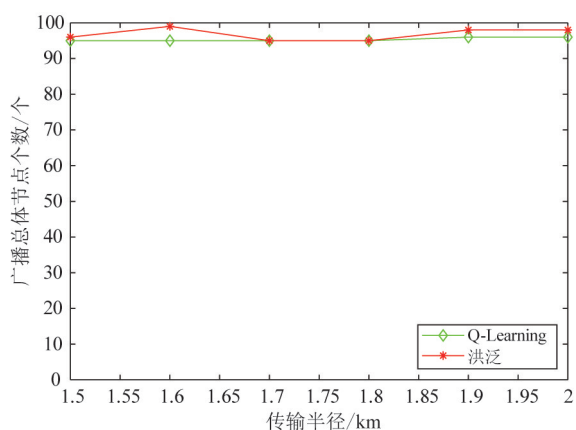


图 5 两种算法关于总体广播节点个数的比较

图 6 显示了两种算法随着节点传输半径的增大,广播过程总能耗发生的变化。从图 6 可知,在广播过程总能耗这一指标上,基于 Q-Learning 的广播路由算法优于传统的洪泛路由算法,另外还可得知,广播过程总能耗随着节点传输半径的增大而增大。

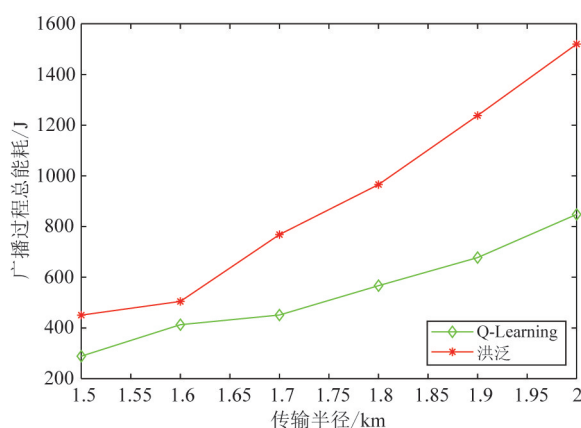


图 6 两种算法关于广播过程总能耗的比较

5 结论

本文对无线区块链网络进行了分析、探讨,并针对无线区块链网络的原理与特点进行分析,提出一

个相适应的无线区块链网络灵活独立的分层逻辑架构。此外,本文分析讨论了现存的无线区块链网络广播路由算法,在常用的洪泛广播路由算法的基础上,针对广播冗余及广播总能耗大等缺陷,提出了一种适用于无线区块链网络的低开销的基于 Q-Learning 的广播路由算法。仿真结果表明,与无线区块链网络中常用的洪泛广播路由算法相比,本文算法有效地降低了广播冗余,其在广播过程总能耗、总体广播节点个数(可达率)、转播节点个数这些性能指标上都有显著的改善效果。

参考文献

- [1] 曾诗钦,霍如,黄韬,等. 区块链技术研究综述:原理、进展与应用[J]. 通信学报, 2020, 41(1): 134-151.
- [2] IANSITI M, LAKHANI K R. The truth about blockchain [J]. Harvard Business Review, 2017, 95(1): 118-127.
- [3] XU H, ZHANG L, LIU Y, et al. RAFT based wireless blockchain networks in the presence of malicious jamming [J]. IEEE Wireless Communication Letters, 2020, 9(6): 817-821.
- [4] 赵瑞琴. 无线多跳网络中若干问题研究[D]. 西安:西安电子科技大学, 2008: 10-13.
- [5] DURRESI A, PARUCHURI V K, IYENGAR S S, et al. Optimized broadcast protocol for sensor networks [J]. IEEE Transactions on Computers, 2005, 54(8): 1013-1024.
- [6] 罗瑛,黄传河,贾小华,等. 一种新的 Ad Hoc 网络中节约能量的广播路由协议[J]. 计算机工程与应用, 2004(35): 154-156, 159.
- [7] SUN Y, ZHANG L, FENG G, et al. Blockchain-enabled wireless Internet of Things: performance analysis and optimal communication node deployment [J]. IEEE Internet of Things Journal, 2019, 6(3): 5791-5802.
- [8] BISWAS A K, DASGUPTA M. Modification of DSDV and secure routing using blockchain technology [C] // 2020 4th International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech). Kolkata: IEEE, 2020: 1-5.
- [9] YU F R, LIU J, HE Y, et al. Virtualization for distributed ledger technology (vDLT) [J]. IEEE Access, 2018, 6: 25019-25028.
- [10] 许子灿,吴荣泉. 基于消息传递的 Paxos 算法研究

- [J]. 计算机工程, 2011, 37(21): 287-290.
- [11] GAO S, YU T. T-PBFT: an eigentrust-based practical Byzantine fault tolerance consensus algorithm[J]. 中国通信, 2019, 16(12): 111-123.
- [12] 夏清, 窦文生, 郭凯文, 等. 区块链共识协议综述[J]. 软件学报, 2021, 32(2): 277-299.
- [13] 李荣, 王芳, 景栋盛, 等. 一种基于 Q 学习的无线传感网络路由方法[J]. 计算技术与自动化, 2017, 36(2): 155-160.
- [14] LI R, FAN L, XIN L, et al. QGrid: Q-learning based routing protocol for vehicular Ad Hoc networks[C]//Performance Computing and Communications Conference. Austin: IEEE, 2014: 1-8.
- [15] WEI T, SUN Y, ZHANG Y, et al. Energy efficient user access and computation offloading strategy for fog radio access network with uplink/downlink decoupling[C]//2019 IEEE 5th International Conference on Computer and Communications (ICCC). Chengdu: IEEE, 2019: 894-900.
- [16] 林艺明. 基于强化学习的车联网可靠路由协议研究[D]. 厦门: 厦门大学, 2017.
- [17] CARTIGNY J, SIMPLOT D, STOJIMENOVIC I. Localized minimum-energy broadcasting in ad-hoc networks[C]//The 22nd Annual Joint Conference of the IEEE Computer and Communications. San Francisco: IEEE, 2003: 2210-2217.
- [18] 唐勇, 周明天. 无线传感器网络中最小化能量广播算法[J]. 通信学报, 2007(4): 80-86.

Wireless blockchain broadcast routing algorithm based on Q-Learning

WEI Jingsheng, SI Pengbo, LI Meng, ZHANG Yanhua

(Department of Information, Beijing University of Technology, Beijing 100124)

Abstract

An independent and flexible layered logic architecture for wireless blockchain network scenarios is proposed with reference to the commonly used blockchain logic layering, and the working principle and role of each layer are introduced in detail. Based on the existing broadcast routing algorithm and combined with the wireless blockchain network scenario, a Q-Learning based broadcast routing algorithm is proposed to overcome the defects of broadcast redundancy and large total energy consumption of broadcast in the broadcast process. The algorithm describes the wireless blockchain network transaction information broadcasting process as a broadcasting decision problem for multiple agents, and helps each current broadcasting node agent to make a decision whether to broadcast or not through offline learning and training. The simulation experimental results prove that the algorithm effectively solves the broadcast redundancy problem and saves the total energy consumption of the broadcast process.

Key words: wireless blockchain, Q-Learning, broadcast redundancy, energy consumption