Engineering Applications of Artificial Intelligence 46 (2015) 23–32

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Application of reinforcement learning to medium access control for wireless sensor networks

Yi Chu *, Selahattin Kosunalp, Paul D. Mitchell, David Grace, Tim Clarke

Communications and Signal Processing Research Group, Department of Electronics, University of York, United Kingdom

ARTICLE INFO

Article history:
Received 6 February 2015
Received in revised form
26 June 2015
Accepted 7 August 2015
Available online 15 September 2015

Keywords:
Reinforcement learning
Wireless sensor networks
Medium access control

ABSTRACT

This paper presents a novel approach to medium access control for single-hop wireless sensor networks. The ALOHA-Q protocol applies Q-Learning to frame based ALOHA as an intelligent slot selection strategy capable of migrating from random access to perfect scheduling. Results show that ALOHA-Q significantly outperforms Slotted ALOHA in terms of energy-efficiency, delay and throughput. It achieves comparable performance to S-MAC and Z-MAC with much lower complexity and overheads. A Markov model is developed to estimate the convergence time of its simple learning process and to validate the simulation results.

1. Introduction

Wireless sensor networks (WSNs) represent a rapidly emerging technology with a wide range of industrial and military applications for environmental monitoring (Akyildiz et al., 2002). A WSN typically comprises a large number of inexpensive nodes, which are capable of sensing, processing and communicating data in a collaborative fashion over multi-hop links such that they are robust to topological changes and failure of a small proportion of the nodes in a network should not affect its operation. Sensor nodes are often battery-powered and when deployed in a distributed fashion in areas which are difficult to access, recharging or replacing the batteries is very difficult. Therefore, energy efficiency is critically important in extending the operational life of such networks and is usually a priority in the design of communication protocols.

In a WSN, nodes share the same medium with their neighbours within a certain range. The Medium access control (MAC) protocol plays an important role in maximising throughput and energy efficiency. A well designed protocol should ensure the successful delivery of the data whilst minimising unnecessary energy consumption arising from collisions and associated retransmissions, control packet overheads, idle listening and overhearing (Demirkol et al., 2006). Many MAC protocols for WSNs have been proposed that significantly improve energy efficiency and throughput performance. However, these improvements incur higher overheads and exhibit ever increasing complexity. Performance evaluation through simulation is common but the practicality of recently proposed schemes is questionable. Given the longer term vision of huge numbers of devices embedded into machines everywhere, much simpler protocols are needed that can nonetheless provide energy-efficient communication, good throughput and acceptable delay.

ZigBee (The ZigBee Alliance, 2015) is a commercial specification for small and low data rate personal area networks, and it is designed based on IEEE 802.15.4 standard (IEEE Standard for Local and Metropolitan area networks, 2012). The MAC layer channel access in IEEE 802.15.4 mainly uses Carrier Sense Multiple Access (CSMA) and ALOHA for random access, and slot scheduling for contention-free access. CSMA based schemes avoid packet collisions by using channel sensing and hand shake procedures (send requests and responses before actual data transmission), however these increase system overheads as well as energy consumption, and still cannot completely eliminate the hidden terminal problem. Slot scheduling ensures success of transmissions but it brings more complexity and computation to the system, while having more overheads than random access schemes.

ALOHA based schemes do benefit from simplicity, low computation and overheads, but they suffer from poor throughput and low energy efficiency through collisions which arise from their blind transmission strategy. Intelligent selection of transmission times offers the potential to significantly improve channel performance and energy efficiency. In this paper, reinforcement learning is applied to ALOHA for this purpose. It enables nodes to develop

* Corresponding author at: Communications and Signal Processing Research Group, Department of Electronics, University of York, United Kingdom.

E-mail addresses: yi.chu@york.ac.uk (Y. Chu), sk772@york.ac.uk (S. Kosunalp), paul.mitchell@york.ac.uk (P.D. Mitchell), david.race@york.ac.uk (D. Grace), tim.clarke@york.ac.uk (T. Clarke).

http://dx.doi.org/10.1016/j.engappai.2015.08.004
0952-1976/

a more effective transmission strategy based on their prior experience. Compared with CSMA based schemes, channel sensing and hand shake procedures are not required, and the only overheads are Acknowledgement (ACK) packets which are commonly implemented in MAC protocols. Nodes can sleep when they have no data to transmit to extend the network lifetime. Compared with contention-free schemes, the learning schemes can achieve the same collision-free channel access without any scheduling information exchange. The only cost is the time for learning algorithm to converge to an optimal steady state, which this paper provides a detailed investigation in Sections 4–6.

Reinforcement learning enables entities to learn effective strategies through trial-and-error interactions in a dynamic environment, with future actions determined by prior experience (Kaelbling et al., 1996). This established artificial intelligence strategy has recently been applied to communication problems and MAC layer protocols. In (Li et al., 2010) and (Tang et al., 2011), several reinforcement learning based MAC protocols are proposed. Although they are not designed for WSNs, a similar strategy can be brought to the design of MAC protocols for WSNs. In this paper, one reinforcement learning method, Q-Learning, is applied to an adaptation of slotted ALOHA as an intelligent slot selection strategy to avoid collisions and retransmissions in single-hop networks. A similar approach has also been evaluated for multi-hop WSNs in (Chu et al., 2012a,2012b). The primary purpose of this paper, here, is to provide a thorough analysis of the application of reinforcement learning to the medium access control problem. A new novel protocol is introduced for single hop networks (ALOHA-Q), which achieves perfect scheduling in steady states with the only overheads being ACK packets. A Markov model of the ALOHA-Q learning scheme is developed to analyse its convergence behaviour and to validate a simulation model. The steady-state performance of ALOHA-Q is then evaluated and compared with S-MAC and Z-MAC through simulation.

The rest of the paper is organised as follows. Section 2 presents related work on MAC protocols for WSNs. Section 3 introduces the principles behind combining Slotted ALOHA and Q-Learning and introduces the ALOHA-Q scheme. Section 4 discusses the tradeoff and fairness issues of the ALOHA-Q scheme. Section 5 provides a Markov model to estimate the convergence time of the learning process. Section 6 presents simulation results, and Section 7 concludes the paper.

2. Related work

A large number of protocols have been proposed for WSNs in an effort to improve energy-efficiency and provide good quality of service in terms of throughput and delay. The majority are contention-based and derived from CSMA since this is a natural approach for distributed multi-hop networks and is employed in the IEEE 802.15.4 standard. Variants of CSMA are also well established for single-hop networks due to standardisation of IEEE 802.11. Some of the more pertinent scheme are summarised in this section.

To improve energy efficiency and the lifetime of a sensor network, S-MAC (Wei et al., 2004) is a well known contention-based scheme which periodically switches nodes between sleep and listening modes. During the listening period, nodes exchange synchronisation and scheduling information and during the sleep period, nodes can either initiate data transmission or remain in sleep mode. The data transmission process is the same as that employed in IEEE 802.11 with Request to Send (RTS) and Clear to Send (CTS) packets exchanged prior to data packet transmission. S-MAC effectively increases network lifetime through the duty cycle, but this serves to increase delay and reduce throughput in densely deployed networks.

Based on S-MAC, DS-MAC (Lin et al., 2004) applies a dynamic duty cycle to achieve a better balance between latency and energy consumption at different traffic levels. The packets in DS-MAC have a field to record their one-hop latency which is used to estimate the current traffic level. Besides the average latency, nodes also keep track of energy consumption on a per packet basis and use this as an indicator of energy efficiency. If the latency of a node rises above a certain level, it will change its duty cycle so that it remains awake for the same duration but reduces its sleep time. It offers better channel performance than S-MAC at high traffic levels, but similar to S-MAC, latency is still a problem.

Demand Wakeup MAC (DW-MAC) (Sun et al., 2008) is another extension of S-MAC which introduces a new low-overhead scheduling algorithm to wake up nodes when they need to transmit or receive in order to increase the effective channel capacity and adapt the network to a variety of traffic loads.

Instead of RTS and CTS packets, DW-MAC transmits scheduling frames at different time points to indicate the length of the data transmissions. It integrates scheduling and access control to achieve low latency and high power efficiency. However, this feature also makes DW-MAC very sensitive to synchronisation errors.

Two intelligent CSMA based schemes are proposed in (Barcelo et al., 2011, Fang et al., 2013). By applying slots and repeating frames to CSMA, the scheme described in (Barcelo et al., 2011) starts with random access, and users continually select slots with successful transmissions until two consecutive collisions. Similarly, the scheme presented in (Fang et al., 2013) is initialised to random access, and the user keeps using the slot with correct packet reception until a collision, then this slots have a decreasing probability of reselection (other slots have equal probability of selection). The embedded intelligence further avoids collisions and improves the channel performance.

By combining the advantages of TDMA and CSMA, Z-MAC (Rhee et al., 2008) achieves better channel utilisation and lower latency. Nodes broadcast ping packets for neighbour discovery after initialisation and select unique time slots according to DRAND (Rhee et al., 2006) to avoid the hidden terminal problem. Nodes have two modes, low contention level (LCL) and high contention level (HCL). At the LCL, any node can contend to transmit in any slot but at the HCL, only the owners of the slot and their one-hop neighbours are allowed to contend for the channel. The owner of the slot always has the highest priority but if it does not have any data to send, other nodes can steal the slot. The priority of the slot owner is applied by using a certain number of contention slots (which are much shorter than data slots) while transmitting packets. The slots owners can always transmit before the non-owners, so that the non-owners can overhear the transmission and avoid collisions. The selection of the length of each contention slot is based on the precision of the synchronisation. Z-MAC performs well at different traffic levels and has robustness to synchronisation errors with the cost of overheads.

Quorum-based MAC (Q-MAC) (Chao and Lee, 2010) adapts sleep schedules to improve energy efficiency and delay performance. The quorum-based wake up scheme determines the wake-up frequency of a node by considering current traffic load. Nodes with a light load switch to a sleep mode more frequently. To reduce the delay caused by long sleep periods, Q-MAC applies a list of next-hop nodes to increase the transmission opportunities for relaying packets.

Spatial Correlation-based Collaborative MAC (CC-MAC) (Vuran and Akyildiz, 2006) improves energy efficiency by making the data transmission process more selective according to the event-based characteristic of WSNs. When an event occurs, nodes within a