TIS: Transmitter Initiated Scheduling for Wireless Sensor Networks

Adrian Udenze
Department of Electronics and Computer Enginering
Nnamdi Azikiwe University Awka, Nigeria
Email: ai.udenze@uizik.edu.ng
Tel: +2348038077638

Abstract: A scheduling based Media Access Control (MAC) protocol named Transmitter Initiated Scheduling (TIS) is presented. TIS maintains multiple schedules tailored to traffic events per communication link resulting in increased performance compared to scheduling strategies with a single schedule for all communication links. TIS is formulated as a queuing system where packets arriving at a node are serviced depending on the node's state. The reward structure for TIS maps directly to network optimisation objectives, namely, minimising average energy consumption subject to average queue lengths. The theoretical framework for TIS is a continuous time Semi Markov Decision Process as opposed to the more frequently used Discrete Time Markov Decision Process. This allows nodes to have multiple sleep states with no fixed frame lengths significantly reducing packet collisions. TIS is suitable for both environments with memory-less event occurrence as well as systems with memory. Simulation results show that TIS offers significant energy savings compared to RL-MAC, a state of the art MAC, for a given delay constraint.

Keywords: Wireless Sensor Networks (WSN); Scheduling; Media Access Control (MAC); Reinforcement Learning (RL); Resource management

1. INTRDUCTION

Consider an event driven network of battery powered nodes: it is unlikely that the nodes will be active, sending and receiving communication packets continuously. Consequently, networks without a scheduling strategy will suffer from power wastage commonly referred to as *Idle Listening* [1, 2]. To prolong node life and ultimately network lifespan, it is desirable for inactive nodes to be powered down or put into lower power states by switching off idle components of a node's system including the radio. Nodes with their radios turned off cannot communicate with neighbouring nodes. Herein lie's the scheduling problem: How long should a node go to sleep for so as to maximise the network lifespan without introducing unacceptable delays to the network? Example applications of Wireless Sensor Networks WSN include: intruder detection, monitoring of animal migration, forest fire detection and environmental monitoring. Events in these types of networks, termed event driven networks [3, 4] are likely to occur after long periods of inactivity or in bursts. Figure 1 below depicts event arrivals in a typical event driven WSN. Short active periods of non deterministic lengths are followed by long idle periods of non deterministic lengths.

Furthermore, events are not always best modelled using memory-less distributions. It may be necessary to use different models for different states of the system. A bursty system being a good example: Work done in [5] shows that in a target tracking task, sensed events exhibited bursty behaviour which was best modelled using Pareto distributions during the *Off* periods and a Normal distribution during the *On* periods. Work done by Simunic et al. [6] also measures the rate of arrivals of service requests for a WLAN card amongst other devices and found that event arrivals were best modelled using multiple distributions. In the active state following an event arrival, events fitted Exponential distributions whereas long idle periods following busy periods was best modelled using a long tail Pareto distribution.

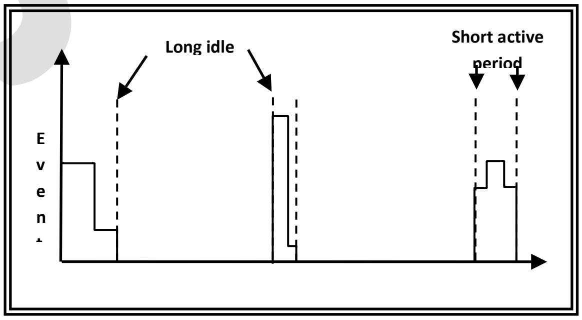

Figure 1 Events in a typical WSN.

The work presented in this paper is a Reinforcement Learning (RL) based MAC technique termed TIS: Transmitter Initiated Scheduling. TIS is light weight, enables nodes to take near optimal actions autonomously in a non deterministic and dynamic network environment. These actions result in schedules that minimise power consumption subject to delay constraints. TIS has a few novel features namely: TIS is formulated as a queuing system and as such is able to adapt to variations in traffic patterns; TIS tailors schedules to each communication link in a network; TIS has multiple sleep states, nodes can sleep for varying lengths of time, this is also well suited to environments with events that fit probability distributions with memory and or multiple distributions; TIS is formulated as a continuing discounted RL problem under a Semi Markov Decision Process (SMDP) framework and as such the rewards structure maps directly to network objectives; nodes also do not have to wake up at fixed discrete points in time which significantly reduces chances of packet collisions; scheduling is based on a low cost TDMA protocol which only requires a field in a header file for control.

The remainder of the paper is organised as follows – a review of related research is presented in section 2. Section 3 describes the TIS protocol and associated mechanisms. Section 4 gives evaluation results followed by conclusions in section 5.

2. Related work

MAC protocols have a number of desirable attributes namely: collision avoidance, energy efficiency, good channel utilization, scalability, low latency, good fairness and throughput as well as good channel utilization [2]. Some of these attributes conflict with each other, the extent of this conflict being dependent on the specific type of network. For example in event driven networks, energy efficiency requires idle radios to be turned off, however due to the non deterministic nature of events, latencies introduced as a result of turned off radios have to be carefully considered. For continuous networks, collision avoidance may result in low utilization of the channel.

MAC protocols broadly fall into two groups, namely: contention and schedule based protocols. Each group has strengths in some of the attributes detailed above at the expense of others. A third group may be thought of as hybrids where the best qualities of the two main groups are merged.

Scheduled MAC protocols are predominantly TDMA based, where competing nodes are given time slots for channel access. Works in [2, 7, 8, 9, 10] all give examples of TDMA based scheduled MAC protocols. The key strength of TDMA based MAC is high energy efficiency and the potential of eliminating collisions. These strengths are however at the expense of limited scalability and adaptability to changes in the network. Also, control is often centralized in order to limit control packet overhead, which leads to hierarchical organization of nodes or cluster formations. The result of which is a limited support of peer to peer communication. A CDMA alternative to TDMA scheduled protocol is discussed in [11] however, CDMA is, due to Multiple Access Interference (MAI), predominantly used in cellular networks.

Contention based MAC protocols do not attempt to eliminate collisions completely but aim to avoid them and recover from them when they occur. The first random access MAC was ALOHA [12], followed by CSMA [13] and other variations such as CSMA/CA [1, 15, 16]. Contention based protocols have increased flexibility and scalability, in addition, nodes can operate on a peer to peer basis. Control overhead is also significantly reduced compared to scheduled protocols. However, the fundamental disadvantage of contention based protocols is energy inefficiency. Collisions may occur, nodes overhear other nodes and idle listening prevention is not implicit. PAMAS [17] attempts to reduce overhearing by introducing duty cycles. A separate signalling channel is used to broadcast a busy signal when the channel is being used thus preventing other nodes from transmitting and also, nodes can turn off their radios for the duration of the transmission thus reducing overhearing.

Hybrid protocols attempt to combine the desirable qualities of both schedule and contention based protocols. A hybrid CDMA / FDMA protocol is presented in the work by Liu et al. [18] in which the authors use FDMA to combat MAI in CDMA protocols. A hybrid CDMA / TDMA protocol is presented by Sohrabi and Pottie, [19]. Nodes maintain a TDMA based schedule for each communication link and CDMA is used to prevent collisions due to neighbouring nodes communicating simultaneously over other links. Bluetooth, [20] is another hybrid CDMA / TDMA protocol where collisions within clusters are avoided using CDMA and communication between cluster heads use TDMA.

Duty cycle based techniques can also be thought of as a hybrid between scheduling and contention based protocols. The general form is for neighbouring nodes to maintain schedules of active and sleep times. During sleep times, nodes power off their radios and go to sleep, in the active period, CSMA/CA is used for media access. The IEEE 802.11 protocol, [16] has a power saving mode in which nodes adopt a duty cycle. Nodes are required to wake up at fixed points in time to listen out for beacons which contain messages to nodes for which messages are queued. If a node receives notice of a message during an Ad-hoc Traffic Indication Message window (ATIM), it stays awake until it receives the message before going back to sleep. Other nodes that do not have pending messages can go back to sleep after the beacon interval. Variations of this technique aim to improve the protocol by reducing the communication overheads incurred by the protocol. Studies in [21 and 22] analyze and discuss impacts of the beacon interval and window sizes on power saved. The IEEE 802.11 protocol also includes MAC in the form of carrier sensing where nodes exchange Request to Send / Clear to Send (RTS/CTS) messages before proceeding with packet transmissions.

An improvement on the IEEE 802.11 protocol is the S-MAC protocol [23], where nodes form virtual clusters around common schedules. By maintaining different schedules for each cluster of nodes, only nodes within a cluster have to wake up at the schedule determined times. This is in contrast to the IEEE 802.11 protocol, where all nodes have to wake up at fixed points in time. Time is divided into frames where each frame consists of an active and sleep phase. In the active phase, the RF circuits are on and a node can communicate with its neighbours. In the sleep phase, the RF circuits are off and a node cannot communicate. T-MAC [24] improves on S-MAC by providing a timeout mechanism in which nodes following a short period of listening for a RTS/CTS exchange can then go to sleep if none is detected. As such, more energy is saved compared to the fixed active and sleep duty cycles of S-MAC. P-MAC [25] improves on T-MAC and S-MAC by tailoring each node's duty cycle to its rate of traffic. Thus nodes with higher rates of traffic have longer duty cycles compared to those with less traffic, thereby improving throughput within the network. RL-MAC [26] is a dynamic protocol that, similar to P-MAC, tailors each node's schedule according to its traffic and that of its neighbours. In contrast to P-MAC however, RL-MAC does not require information of traffic flow in advance of scheduling. The RL-MAC protocol learns by interacting within the network environment, near optimal duty cycles which makes it able to adapt to dynamic events in a network. RL-MAC assumes a traffic pattern that follows an exponential distribution and maintains one schedule for all communicating links.

TIS i designed to improve on RL-MAC by being suited to event driven networks with bursty complex traffic patterns. Figure 1 shows a star topology in which nodes 1-6 communicate with a sink node. Such a topology models a network in which multiple nodes are trying to transmit messages to a single node. Figure 2 shows a duty cycle timeline for RL-MAC and TIS under bursty network conditions. RL-MAC due to its single schedule has to stay awake for long periods to service events from all nodes whereas TIS due to its multiple schedules takes better advantage of long idle periods resulting in increased energy savings.

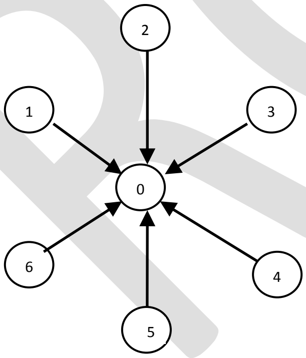

Figure 2: A star topology

3. TIS

The fundamental idea behind TIS is that all nodes where traffic originates, termed *source nodes*, are responsible for directing their own schedules. Source nodes may be leaf nodes or cluster heads in a clustered WSN. TIS operates under a SMDP framework and as such nodes can sleep for any length of time and stay awake for any length of time. Also, TIS has multiple sleep states enabling nodes to sleep for longer during long idle periods. A field is added to the data packet header called SCHEDULE, the length of which is dependent on the set of actions available to a node. Starting at the beginning of the cycle of events, all nodes are on. Following an appropriate routing protocol, during which communication paths are established, each node is able to communicate with its neighbours or cluster heads. Schedule information contained in the SCHEDULE field is contained in each communication. The information contained informs receiving nodes of the senders schedule comprising, proposed sleep time durations i.e. time of next transmission following sleep state transitions. The receiving node replies the transmitting node (initiator) with confirmation of the proposed schedule. The schedule is thus fixed for the next communication over the specific link and the same process establishes schedules for all communication links. The SCHEDULE field and the information contained can be thought of as similar to the patterns field in P-MAC. The schedules formed are thus dynamic, dependent on each node's traffic and only nodes expecting communication need wake up from inactive modes. The only condition imposed on the sleep time is that it is greater than the break even time of the node. The break even time is the minimum amount of time a node must sleep to ensure energy savings [6]. If no RTS signal arrives during a short time out period, the receiving node assumes no communication is pending and transitions to the next sleep state for that link. Note that the continuous time nature of TIS means that nodes do not all wake up at fixed times, unlike RL-MAC and S-MAC and as such the chances of media contention is significantly reduced. Where schedules do clash and more than 1 node is awake to transmit to a receiver at the same time, CSMA/CA is used for collision avoidance. The resulting delay is also reflected in the reward structure for actions taken such that conflicting nodes will take actions to resolve the conflict at the next decision epoch. Each schedule has an associated cost in terms of queue sizes and energy consumed. Every time an event is generated, it is time stamped. Upon waking up from a sleep cycle, communication links are on. The packets received contain time stamps of the time of generation which is used to calculate the cost of the sleep cycle. In this sense, costs are delayed one cycle of events. Furthermore, schedules for time $t+1$ are proposed at time t, and finally all actions are proposed in the active state. Multiple sleep states exist: If a node wakes up from an inactive state and no events are queued, it will transition to the next sleep state, where a decision is made on the length of sleep. After a due learning process, all nodes reach optimal actions. The following notes are made or reiterated: (1) to minimise transmissions, receiving nodes may assume no communication is pending and go to sleep after an appropriate set time (the set time is assumed to be significantly greater than the clock drift to eliminate synchronisation loss). The next transmission time will be as previously scheduled. (2) Times are relative to communications and not absolute, as such loss of synchronisation due to clock drift is reduced, [23]. (3) To accommodate new nodes joining the network and recovery from loss of synchronisation, nodes may be required to wake up at fixed points in time, as in IEEE 802.11 MAC protocol. The time between beacon intervals and the ATIM window is left to the discretion of the network designer / application.

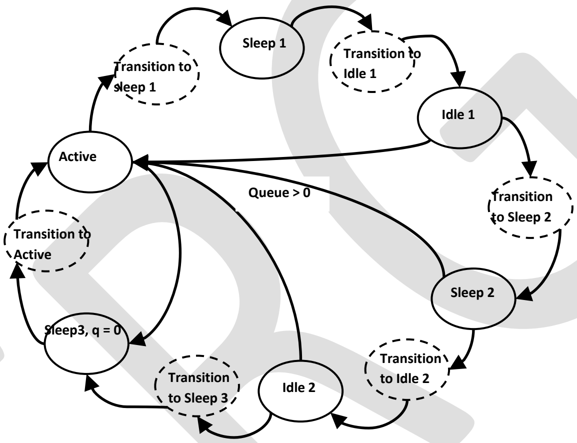

Figure 4: Node state transition diagram showing multiple sleep and idle states

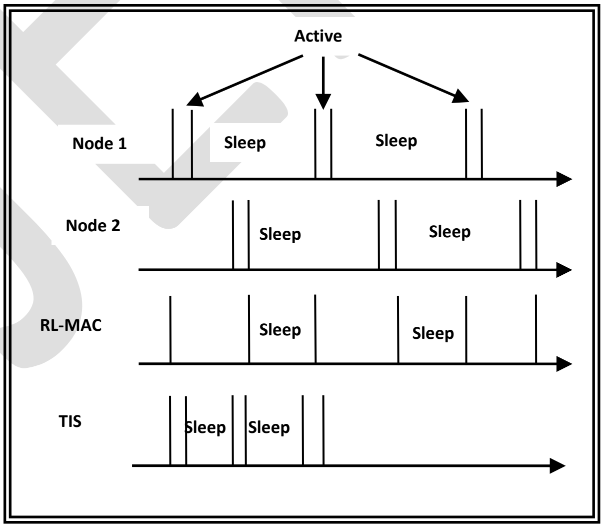

Figure 3: RL-MAC Vs. TIS duty cycle timeline

In summary, the proposed protocol ensures all communicating nodes know when to expect packets. The child nodes or source nodes are responsible for proposing schedules to parent or receiving nodes.

Theorem (1): The optimal scheduling policy can be found by solving:

$$h(s) = \min_{a \in A} \left\{ \mathrm{cost}(s,a) - g(s)y(s,a) + \sum_{j \in S} m(j \mid s,a)h(j) \right\} \qquad (1)$$

Where *h(s)* is referred to as the bias, *g(s)* is the average cost and *m(j|s,a)* is as defined above.
Proof of the theorem and an in-depth study of SMDPs can be found in [27 and 28].

3.1 R-Learning

The problem defined in (1) can be solved using several techniques including Value Iteration, Policy Iteration and Linear Programming (LP) where models of the environment are available, [27]. Reinforcement Learning (RL) on the other hand offers solutions by interacting directly with the environment, observing rewards and subsequently learning optimal behaviour. [29] offers a comprehensive introduction to RL techniques. The problem faced in a WSN scheduling problem as formulated in this work, is to minimise energy used up by nodes subject to performance constraints over the lifetime of the network. To this end the TIS agent's rewards are mapped directly to the optimisation objectives, namely energy consumed and the length of the queue at each time step for the life of the WSN node. Furthermore, the energy used up and the queue length in the present is worth exactly the same as in the future and as such, these are best treated as undiscounted problems. To this end R-Learning, a form of RL for undiscounted continuing tasks is preferred over Q learning.
Value functions are defined relative to the average expected reward per time step under a given policy π:

$$\rho^{\pi} = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} E_{\pi} \{ r_t \} \qquad (2)$$

The process is assumed to be ergodic, consequently ρ^{π} is independent of a starting state. The value of a state is defined as:

$$\tilde{V}^{\pi}(s) = \sum_{k=1}^{\infty} E_{\pi} \{ r_{t+k} - \rho^{\pi} \mid s_t = s \} \qquad (3)$$

and the state action value function defined as:

$$\tilde{Q}^{\pi}(s,a) = \sum_{k=1}^{\infty} E_{\pi} \{ r_{t+k} - \rho^{\pi} \mid s_t = s, a_t = a \} \qquad (4)$$

n and *k* denote decision epochs in this case (a SMDP problem), $\tilde{Q}^{\pi}(s,a)$ and $\tilde{V}^{\pi}$ are relative values to the average reward under the current policy. In state $a \in A_s$ the agent chooses action $a \in A_s$ using a suitable behaviour policy, i.e. ϵ-greedy [29] and receives an immediate lump sum reward *l(s,a)*. The lump sum reward is the cost of making a transition. A consequence of choosing action $a \in A_s$ is that the next decision epoch occurs *t* time units later during which time the system may transition between states at a cost $c(j', s, a)$. The cost $c(j', s, a)$ is the energy consumed and the queue length for the time spent in each state between decision epochs. The full R-Learning algorithm as employed in this work is given below:

```
Initialize $\rho$ and $Q(s,a)$, for all $s$, $a$, arbitrarily

Repeat forever:

        $s$ ← current state

        Choose action $a$ in $s$ using ϵ-greedy
        behaviour policy

        *Take action $a$, observe immediate $r$
        and subsequent delayed $r$ in $s'$

        Repeat for each time step that
        occurred between epochs:

        $Q(s,a) \leftarrow Q(s,a) + \alpha[r - \rho + \max_{a'} Q(s,a)$
        $- Q(s,a)]$
```

Table 1: R-learning algorithm for the TIS agent

3.2 TIS Rewards

The choice of actions is the length of time to spend in each sleep state i.e. 1s, 2s, …10s.. The cost of each action is the energy consumed and the length of the queue at each time step for the duration of time spent in each state. Also, rewards are delayed until the system transitions to an active state, when a node knows how long the queue is, having received queued packets from communicating nodes.

3.3 Multi Criteria RL

The TIS problem has two conflicting objectives namely: to reduce queue length or delay and at the same time to reduce energy consumption. Energy reduction results in increased delays and vice versa.

This class of RL problems is known as multi-criteria problems. One approach to determining optimal actions would be to use an ordering of returned rewards. In the case of the work presented, assuming a problem setup of minimising average energy consumption subject to a fixed average queue length constraint, then optimal actions could be chosen as follows. Keep running averages of both queue lengths and energy consumption. Select all actions that lead to an average queue less than the fixed constraint. Among the selected actions, the optimal action is the one with the least energy consumption. This forms the basis for the work presented by Gabor et al. [30]. Another alternative is to convert the vector reward function to a scalar which can then be treated as a standard RL problem. This conversion from a vector to a scalar function is done using weights, [31]. Weights are chosen to reflect the levels of priority assigned to each reward value and the sum of all weights equals 1. For example, a weight of 0.5 assigned to energy consumed and 0.5 assigned to queue lengths denotes equal priority to both rewards.

To ensure no bias is introduced by either reward value, the data is first normalised. For the purpose of the work done here, normalisation was carried out using:

$$Value' = \frac{Value - OriginalMin}{OriginalMax - OriginalMin}(NewMax - NewMin) + NewMin \quad (5)$$

4 Simulation results

Bursty arrival - Exponential distribution	
Parameter	Value
λ (average)	2 per min
Bursty arrival duration – Normal distribution	
Parameter	value
μ (mean)	5 min
σ^2 (variance)	1 min
Long idle periods – Tail Pareto 1 – at^{-b}	
a	0.9
b	0.01
Node service times – Exponential distribution	
Parameter	Value
λ (average)	10 per sec

Table 2: Parameters of models fitted to measured data

Parameter	Value
Max Queue length	20
Action Value LR	α = 0.1
Average Value LR	β = 0.01
Transmission power	36mW
Receiving power	14.4mW
Sleep power	0.015mW
Transition power	28.8mW

Transition time	0.8ms

Table 3: Simulation parameters

RL-MAC is a scheduling technique that improves on T-MAC, P-MAC and S-MAC and as such is used as a bench mark for TIS. A simulation environment was set up using the above models. The star topology of figure 1 was simulated to examine the performance of RL-MAC against TIC when multiple nodes try to communicate to a sink.

Figure 10 below shows a comparison of average rewards received for RL-MAC and TIS. The rewards are a vector conversion of queue lengths and energy per time step to a scalar reward value as described in section 3.4. TIS 2 shows the results of a two sleep state implementation where depending on whether or not events are queued during the first sleep state, a node goes into a longer sleep state where more energy can be saved. TIS 3 shows a three sleep state implementation where nodes can go into a third sleep state for longer sleep as shown in figure 4 above. TIS 3 receives the most rewards followed by TIS 2, both TIS implementations perform better than RL-MAC.

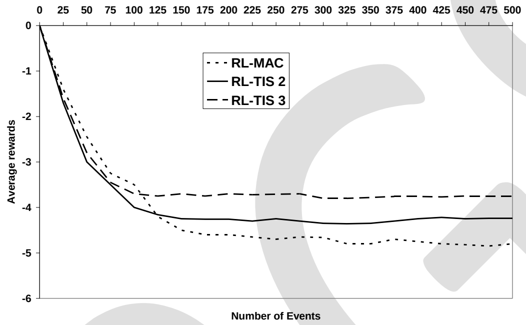

Figure 5: Reward comparison of RL-MAC to two and three sleep state versions of TIS

The next set of results show explicit results of average delay and average energy consumption for RL-MAC and two versions of TIS as described above. For the delay costs, each time an event occurs, the time of arrival is recorded and for every time step until the event is serviced i.e. reaches intended destination, the delay is incremented. The energy costs are the power used up per time step in each state. Value functions were maintained for both competing objectives and the best actions were chosen in an ordered manner as described in section 3.4. It is assumed that the frame sizes and therefore duty cycles in the RL-MAC protocol can be tuned to traffic in network nodes as such a fair comparison should show improved energy figures for similar levels of delay. The results in figures 11 and 12 show just this. For a slight increase in average delay, figure 11 (less than 4% increase) a decrease in energy consumption of more than 20% was observed for TIS 3, figure 12.

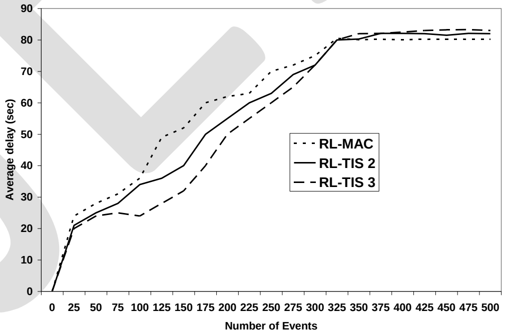

Figure 6: Average delay comparison of RL-MAC to two and three sleep state versions of TIS

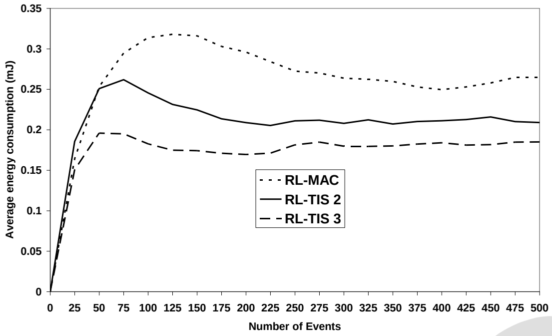

Figure 7: Average energy comparison RL-MAC to TIS

Conclusions

Traffic patterns in event driven WSNs do not always follow a memory less distribution. For such applications TIS introduces multiple sleep states and multiple schedules to take optimal advantage of long idle periods. TIS is formulated as a continuous RL task under a SMDP framework, such that nodes do not have to wake up at fixed discrete times. TIS is truly autonomous, dynamic and able to operate in a clustered or peer to peer network with limited control overheads. Simulation results show that TIS outperforms RL-MAC a state of the art protocol in terms of energy consumption for a given delay constraint.

Acknowledgements

This research is supported by Udala research laboratories and Ezned Nigeria Limited.

REFERENCES:

[1] Udenze, A. (2014), “HYMAC: An intelligent collision avoiding MAC for Wireless Sensor Networks”, in the International Journal of Engineering Research & Technology, Vol 3, Issue 11.
[2] Sichitiu, M. L. (2004) ‘Cross-layer scheduling for power efficiency in wireless sensor networks’ In IEEE Infocom. 3, pp. 1740-1750.
[3] Wang, P. and Akyildiz, I. F. (2009) “Spatial correlation and mobility aware traffic modelling for wireless sensor networks’, in Proc. of IEEE Global Communications Conference (GLOBECOM’09.
[4] Wang, Q. and Zhang, T.(2008) “Source traffic modeling in wireless sensor networks for target tracking”, in Proc. of the 5th ACM International Symposium on Performance Evaluation of Wireless Ad-Hoc, Sensor, and Ubiquitous Networks (PEWASUN’08), pp. 96–100.
[5] Qinghua, W. and Tingting, Z. (2008), ‘Source Traffic Modeling in Wireless Sensor Networks for Target Tracking’, Int. Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems, British Columbia, Canada, pp. 96-100.
[6] Simunic, T., Benini, L.,Glynn, P. and Giovanni De Micheli, G. (2001) ‘Event driven power management’ IEEE Trans. On CAD of IC and Sys. 20, pp. 840-857.
[7] Hohlt, B., Doherty, L. and Brewer, E. (2004) ‘Flexible power scheduling for sensor networks’ In Proceedings of the third international symposium on Information processing in sensor networks, pp.205 – 214.
[8] Hohlt, B., and Brewer, A. (2006) ‘Twinkle: Network power scheduling for tinyos applications” In IEEE international conference on distributed computing sensor systems, San Francisco, USA.
[9] Arisha, K. A., Youssef, M. A. and Younis, M. F. (2002) ‘Energy-aware TDMA based MAC for sensor Networks’, Proceedings of IEEE IMPACCT 2002, NY, USA.
[10] Heinzelman, W. R., Chandrakasan A. and Balakrishnan, H. (2000) ‘Energy-efficient communication protocols for wireless microsensor networks’, In proceedings of the Hawaii International Conference on Systems Sciences.
[11] Proakis, J. G., Salehi, M. and Bauch, G. (2003) ‘Contemporary Communication Systems using MATLAB and Simulink’ 2nd ed. Southbank, Vic. Thomson/Brooks/ Cole.
[12] Abramson, N. (1970) ‘The ALOHA system – Another alternative for computer communications’, Proceedings of the November 17-19, 1970, Joint Computer Conference Houston, Texas, pp. 281-285.
[13] Kleinrock, L. and Fouad, T. (1975) ‘Random access techniques for data transmission over packet-switched radio channels’ Proceedings of the national computer conference and exposition, Anaheim, California, pp. 187-201, May 19-22.
[14] Karn, P. (1990) ‘MACA - A new channel access method for packet radio’ In. Proceedings of Amateur Radio Ninth Computer Networking Conference, ARRL, pp. 134-140.

[15] Bharghavan, V., Demers, A., Shenker, S. and Zhang, L. (1994) 'MACAW: A media access protocol for wireless LANs', Proceedings of the conference on Communications architectures, protocols and applications**,** London, pp. 212 – 225.

[16] LAN WAN Standards Committee of the IEEE Computer Society (1997) 'Wireless LAN medium access control (MAC) and physical layer (PHY) specification' IEEE, New York, NY, USA, IEEE Std 802.11-1997.

[17] Singh, S. and C. S. Raghavendra, C. S. (1998) 'PAMAS: Power aware multi-access protocol with signalling for ad hoc networks', ACM Computer Communications Review, vol. 28, no. 3, pp. 5–26, July.

[18] Liu, B. H., Chou, C. T., Lipman, J. and Jha, S (2005) 'Using frequency division to reduce MAI in DS-CDMA wireless sensor networks', Wireless Communications and Networking Conference, IEEE Vol. 2, 13-17 March, pp. 657 – 663.

[19] Sohrabi, K. and Pottie, G. J. (1999) 'Performance of a novel self-organizing protocol for wireless ad hoc sensor networks', In proceedings of the IEEE 50th Vehicular Technology Conference, pp 1222-1226.

[20] Bluetooth SIG Inc., (2001) 'Specification of the Bluetooth system: Core', http://www.bluetooth.org/.

[21] Jung, E. S. and Vaidyan N. (2002) 'Energy efficient MAC protocol for Wireless LANs. In proceedings of IEEE Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM, 3, pp. 1756 – 1764.

[22] Woesner, H., Ebert, J. P., Schlager, M. and Wolisz, A. (1998) 'Power-Saving Mechanisms in Emerging Standards for Wireless LANs: The MAC Level Perspective' IEEE Personal Communications, Vol. 5, issue 3, pp.40-48.

[23] Ye, W., Heidemann, J. and Estrin, D. (2004) 'Medium access control with coordinated adaptive sleeping for wireless sensor networks', IEEE/ACM Transactions on Networks, Vol. 12, No. 3, pp.493–506.

[24] Dam, T.V. and Langendoen, K. (2003) 'An adaptive energy efficient MAC protocol for wireless sensor networks', SenSys '03: Proceedings of the first International Conference on Embedded Networked Sensor Systems, New York, NY: ACM Press, pp.171–180.

[25] Zheng, T., Radhakrishnan, S. and Sarangan, V. (2005) 'Pmac: an adaptive energy-efficient MAC protocol for wireless sensor networks', Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05), pp.65–72.

[26] Zhenzhen, L. and Itamar, E. (2006), 'RL-MAC: a reinforcement learning based MAC protocol for wireless sensor networks', International Journal of Sensor Networks**,** Vol. 1, Issue 3/4 September 2006, pp. 117-124

[27] Puterman, M. L., (1994) Markov Decision Processes, Wiley Series in Probability and Mathematical Statistics, J. Wiley and Sons.

[28] Udenze A. (2015), "SMDP-MAC: Event driven Media Access Control for WSN", International Journal in IT and Engineering, Vol. 3, Issue 6, June.

[29] Sutton, R. S. and Barto, A. G. (1998) Reinforcement Learning: An Introduction, Cambridge MA: MIT Press.

[30] Gabor, Z., Kalmar, Z. and Szepesvari, C. (1998) 'Multi-criteria Reinforcement Learning', Proceedings of the Fifteenth International Conference on Machine Learning, p.197-205.

[31] Natarajan, S. and Tadepalli, P (2005) **'**Dynamic preferences in multi-criteria reinforcement learning'**,** Proceedings of the 22nd international conference on Machine learning, Bonn, Germany, pp. 601 – 608