Adaptive Wireless Networks Using Learning Automata

P. Nicopolitidis, ***Member, IEEE*****, G. I. Papadimitriou,** ***Senior Member, IEEE*****,**

P. Sarigiannidis, ***Member, IEEE*****, Mohammad S. Obaidat,** ***Fellow, IEEE*** **and A. S. Pomportsis,** ***Member, IEEE***

Corresponding Author: Mohammad S. Obaidat at: obaidat@monmouth.edu

Abstract

Wireless networks are characterized by the fact that they operate in environments with unknown and time-varying characteristics. The changing nature of many of these characteristics will significantly affect network performance, with the most important challenges arising from variable link qualities, dynamic topologies, power management, spectrum allocation and traffic patterns. This fact has a profound impact on the design of efficient protocols for wireless networks and as a result adaptivity arises as one of the most important properties of these protocols. Learning Automata are Artificial Intelligence tools that have been used in many areas where adaptivity to the characteristics of the wireless environment can result in a significant increase in network performance. This paper reviews state of the art approaches in using Learning Automata to provide adaptivity to wireless networking. After discussing the way these tools operate, we discuss state of the art research on using Learning Automata on several layers of the wireless networking stack and on specific wireless networking environments, such as sensor networks and wireless data broadcasting.

Index Terms: Learning automata, wireless networks, adaptive power transmission, adaptive routing algorithms, sensor networks.

1. Introduction

Wireless networks operate within a dynamic environment, which entails possibly unknown and time-varying characteristics, with the most common ones being the following:

Variable link qualities. These are primarily caused by multipath fading and co-channel interference, which result in having channel Bit Error Rates (BER) in the order of magnitude of up to ten times that of a cable's BER. The time variability in the quality of a link leads to the need for adaptive operation of several protocols across the protocol stack. For example, at the physical layer, an increased BER should be countermeasured by either using a more robust modulation scheme (and thus a lower transmission rate), or an increase in transmission power via power control procedures. At the transport layer, the TCP congestion window (cwnd) should be adaptively handled so as not to misinterpret transmission losses for congestion.

Dynamic topologies. Generally, neither a fully connected topology nor a static one between the nodes of a wireless network can be assumed. This is due to the fact that these nodes are mobile and have a fixed transmission range. Thus, the topology of the network will change with time and the network nodes should adapt to such changes. Typical examples of protocols that need to adapt to dynamic topologies are wireless Medium Access Control (MAC) and routing protocols.

Power management. The nodes of a wireless network are usually also mobile and thus are typically battery-powered. Therefore, specific measures have to be taken in the direction of adapting the energy consumption of a node to its residual node energy. Power control mechanisms that are typically employed in wireless networks must also take this need into account.

Spectrum usage. For the past several years, we have witnessed an increasing research efforts in wireless cognitive systems that can opportunistically use licensed parts of the spectrum when these are not in use. Thus, network nodes should adapt their transmission frequency according to the operation of collocated licensed systems, with profound challenges for several protocol layers.

Changing traffic patterns. The needs of nodes for medium access can change over time according to the needs of the applications being served, thus leading to the need for adaptive channel access protocols.

Learning Automata (LA) [1] have been found to be useful in systems which possess incomplete knowledge regarding the environment in which they operate. In the area of data networking, LA have been applied to several problems, including the design of self-adaptive MAC, routing and transport layer protocols. Such protocols have been proposed both for wired and wireless platforms and efficiently operate in networks with dynamic workloads. This paper surveys state of the art approaches in using LA to enhance the performance of wireless networks. The article is organized as follows. After providing an introductory coverage of LA, we discuss state of the art research on using LA on several layers of the wireless networking stack and on specific wireless networking environments, such as sensor networks and wireless data broadcasting networks.

2. Reinforcement Learning and Learning Automata

In the area of Learning systems, Reinforcement Learning has emerged a promising technique. Reinforcement Learning aims to provide a system with the necessary information in order to plan its actions so as to maximize the reward it receives from the environment in which it operates. Reinforcement Learning techniques typically utilize the following triple: i) the number of environmental states S, ii) the number of possible actions M to be taken by the system and iii) an environmental reward β for each action taken by the operating system. At each time instant during its operation, a system based on Reinforcement Learning that resides in a certain state, chooses one of the available actions, performs it and receives new state from the environment as well as the environmental response. By repeating the above procedure, the goal of Reinforcement Learning system is to achieve maximization of the received environmental reward.

Learning Automata are Artificial Intelligence tools whose operation can be viewed in the framework of Reinforcement Learning. The operation of a LA is shown schematically in Figure 1. A LA continuously interacts with the random operating environment so as to find among a set of actions the one that minimizes the average penalty received by the environment. To achieve this, a LA uses a vector $p(n)$, which maintains the probability distribution for choosing at cycle n, the action $\alpha(n)$ from the set of actions $\alpha_1, \alpha_2 \ldots\ldots \alpha_M$. Obviously, $\sum_{i=1}^{M} p_i(n) = 1$. A LA can be used only if there exists a feedback mechanism that conveys to the LA the environmental response to each performed action.

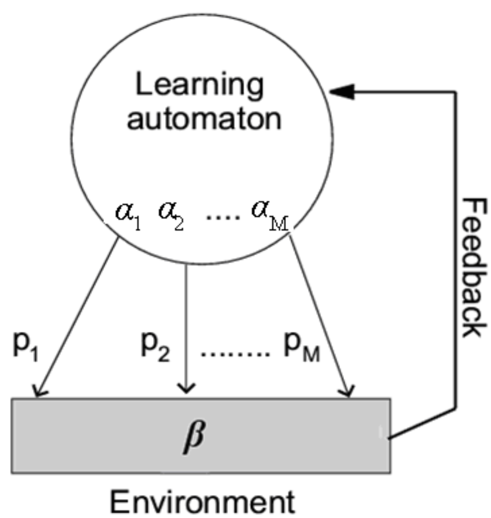

Figure 1: Operation of a Learning Automaton.

The operation of a LA is based on the probability updating algorithm, also known as the reinforcement scheme. This algorithm uses the environmental response that was received as a result of performing the action a_i selected at cycle n (action $\alpha(n)$), in order to update the probability distribution vector p. After the updating is performed, the LA selects the action to perform at cycle n+1, according to the updated probability distribution vector $p(n+1)$. A general reinforcement scheme has the form of Equation (1):

$$
\begin{aligned}
p_i(n+1) &= p_i(n) - (1-\beta(n))g_i(p(n)) + \beta(n)h_i(p(n)), \ \text{ if } \alpha(\mathrm{n}) \neq \alpha_{\mathrm{i}} \\
p_i(n+1) &= p_i(n) + (1-\beta(n))\sum_{j\neq i} g_j(p(n)) - \beta(n)\sum_{j\neq i} h_j(p(n)), \ \text{ if } \alpha(\mathrm{n}) = \alpha_{\mathrm{i}}
\end{aligned} \tag{1}
$$

The functions g_i and h_i are associated with reward and penalty for the selected action a_i, respectively, while $\beta(n)$ is a parameter expressing the received environmental response at cycle n, normalized in the interval [0,1]. The lower the value of $\beta(n)$, the more favourable the response is.

If the environmental response is of a binary nature, indicating only reward or penalty via 0 or 1 respectively, the environment is known as P-model and the LA is known to operate within a P-model environment. Nevertheless, as in many cases a P-model LA will yield only a gross estimation of the environmental response, other schemes have also appeared. In these, the environmental response can take more than one value in [0..1] thus indicating actions that are neither completely rewarding or penalizing. Specifically, in a Q-model LA, the environmental response can take more than two, still finite however, possible values in the interval [0..1], whereas in an S-model LA, the environmental response can take continuous values in [0..1].

Regarding the choices made for functions g_i and h_i, different selections for these functions result in a number of different reinforcement schemes, with the most common ones being the following:

The Linear Reward–Penalty (LR–P) scheme. In this scheme, after selection of an action a_i at cycle n, the following reinforcement scheme is applied:

$$p_i(n+1) = p_i(n) + \beta(n)\ [L / (M\text{-}1) - L\ (p_i(n)\text{-}a)] - [\ 1 - \beta(n)\ ]\ L\ (p_i(n)\text{-}a), \forall \quad \alpha(n) \neq \alpha_i$$

(2)

$$p_i(n+1) = p_i(n) - \beta(n)\ L\ (p_i(n)\text{-}a) + [1 - \beta(n)]\ L\ [1 - (\ p_i(n)\text{-}a)], \text{ if } \alpha(n) = \alpha_i$$

In the above equation pair, M is the number of actions and L is a parameter that lies in (0..1) and defines the learning rate of the LA. In this scheme, g_i and h_i are linear functions of the corresponding action probabilities p_i. Thus, in a P-model LR–P (P_{LR-P}) automaton, after reception of a favourable response for the action that was selected in the previous cycle, the corresponding probability of selecting this action again is increased. After reception of an unfavourable response however, the probability of selecting this action again is decreased. For the Q and S model LR–P schemes probability updates occur as combinations of g_i and h_i weighed by 1- $\beta(n)$ and $\beta(n)$, respectively.

The Linear Reward–Inaction (LR–I) scheme. In this scheme, after selection of an action a_i at cycle n, the following reinforcement scheme is applied:

$$p_i(n+1) = p_i(n) - L(1-\beta(n))(p_i(n)-a) \;\; \forall \; \alpha(n) \neq \alpha_i$$
$$p_i(n+1) = p_i(n) + L(1-\beta(n))\sum_{j\neq i}(p_j(n)-a) \;\; \text{if } \alpha(n) = \alpha_i \tag{3}$$

In this scheme, g_i is a linear function of p_i and h_i always equals 0. Thus, in a P-model LR–I ($P_{LR\text{–}I}$) automaton, after reception of a favourable response ($\beta(n)$=0) for the action that was selected in the previous cycle, the corresponding probability of selecting this action again is increased. When an unfavourable response ($\beta(n)$=1) is received, however, the probability of selecting this action again is not decreased, but remains the same. For the Q and S model LR–I schemes probability updates occur as functions of g_i's only.

Nonlinear schemes. In this case the functions g_i and h_i are nonlinear functions of p_i.

For the parameters L and a in Equations (2) and (3), it holds that L and a take values in (0,1) and p_i takes values in (a,1). L defines the speed of the automaton convergence. The lower the value of L the more accurate the estimation made by the automaton, a choice however that comes at expense over convergence speed. Parameter a is used to enhance the adaptivity of the LA. This is because when the choice probability of an action approaches zero, then this action is not selected for a long period of time. However, after a time period the changing nature of the environment might render this action to be a favorable one. However, since this action now has a zero probability of being selected the LA will not be capable of selecting it any more. Thus, the use of a non-zero value for parameter a prevents the choice probabilities of actions from taking values of zero and increases the adaptivity of the LA by preventing the above phenomenon. When the environment is slowly switching, a and L must be very close to zero in order to guarantee a high accuracy, whereas in a rapidly switching environment higher values of a and L can be used, in order to increase adaptivity.

To better understand parameters L and a, Figures 2 and 3 show their effect on the convergence of a $P_{LR\text{-}P}$ LA. In this experiment, we assume an action α_i with an actual probability d_i=0.8 of being the optimal one, with this probability changing to a new one (dashed line, d_i=0.8) after about 30,000 cycles and we plot the probability estimate of the LA for this action. One can see that the choice for the value of L reflects the classic speed versus accuracy problem. As can be seen from Figure 2, where L=3*10^{-2}, large values of L (compared to that used in Figure 3) provide a higher convergence speed at the expense, however, of convergence accuracy. Contrary, with a small value of L (Figure 3), we get better convergence, at the expense, however, of convergence speed. Finally, convergence accuracy also depends on the value of a, with smaller values of a giving better convergence. Increased values of a will make the estimated probability p_i converge to a point higher than d_i. Figure 3 supports this fact.

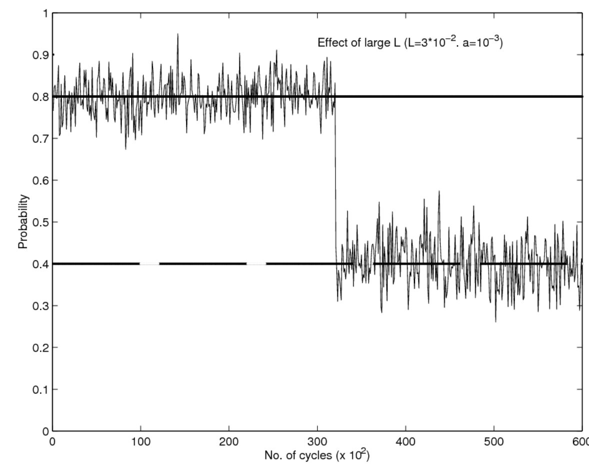

Figure 2. Effects of a large value of the learning rate L on P_{LR-P} LA convergence.

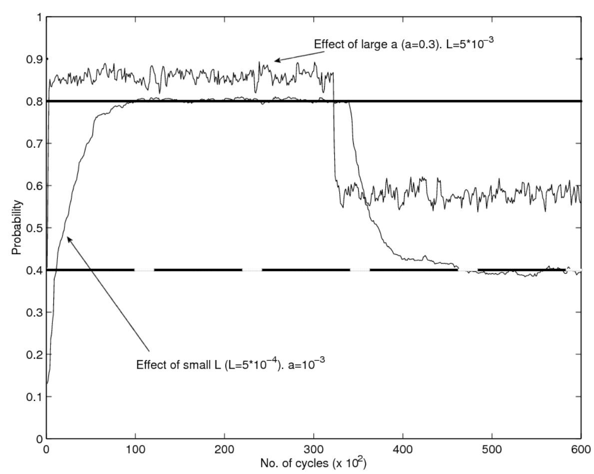

Figure 3. Effects of a small value of the learning rate L and of a large value of parameter a on P_{LR-P} LA convergence

By altering the selection procedure for the action to be rewarded or penalized after each cycle, more efficient LA in terms of convergence speed can be derived. This idea is implemented in pursuit of LA, which employs a vector that contains the running estimates of the environmental response for each action and at each probability update the LA will always reward the action with the current minimum penalty estimate, thus always pursuing the action with the highest reward.

3. Applications of LA to wireless networking

3.1 Applications to different layers of the network

Physical layer

Learning automata have recently been applied to determine the transmission power of mobile nodes. In [2], the authors model the distributed power control problem in an infrastructure wireless network as a non-zero sum game between the mobile nodes. Two LA-based algorithms for distributive solution of the power control game are proposed. In both of them, each node operates a LA, which determines the probability of choosing a certain transmission power based on the feedback received from the Base Station (BS). This feedback is essentially the environmental response for choosing a certain power level and expresses the amount of information that a mobile node can transfer during the lifetime of its battery. However, the proposed algorithms can also operate for other definitions of the environmental response as well.

Learning automata have also been applied to dynamically adapt the transmission rate in wireless links in a physical/MAC layer synergy. In [3], the authors present a joint user scheduling and adaptive rate control for downlink wireless transmission that can be easily applied to Third Generation (3G) wireless systems. User scheduling is performed at the MAC layer and is enabled via a LR-P LA while the rate selection procedure is implemented in the physical layer via a pursuit reward-inaction LA. The respective components of the above two procedures exchange information so as to enable the requests of each user in terms of throughput to be served under the variable transmission rate environment. The joint algorithm is shown to be able to learn and use the best transmission rate according to the conditions of the channel. The learning ability of the channel is enabled by knowledge gained via the acknowledgment packets that return from the receivers of downlink packets. Simulation results show that the proposed algorithm is suitable for low mobility applications in 3G wireless networks.

In [4], a LA-based scheme named Stochastic Automata Rate Adaptation Algorithm (SARA) is proposed in order to achieve rate adaptation. According to this scheme, each transmitter has a list of possible modulation schemes to use, each one yielding a different transmission rate. The selection for the rates to use and the probability of using each one of these rates is dynamically updated via a LA based on the obtained feedback (positive – negative acknowledgements) from the receiver. SARA, which is fully compatible with the IEEE 802.11 MAC protocol, is compared via simulation under different channel scenarios to Automatic Rate Fallback, Adaptive Automatic Rate Fallback and the scheme presented in [3]. The results show superior performance of SARA compared to the previous approaches. Moreover, the results indicated a lower computational complexity for achieving rate adaptation of SARA when compared to the scheme presented in [3].

MAC layer

In the context of medium access, Learning automata have found use in both infrastructure and ad-hoc wireless LANs (WLANs). In [5] and [6] the ability of LA to learn the parameters of the operating environment is exploited in order to provide efficient MAC protocols for bursty-traffic WLANs. It is proved for both the approaches in [5] and [6] that the learning algorithm asymptotically tends to assign to each node a portion of the bandwidth proportional to the node's needs.

In [5], the BS is equipped with a P_{LR-P} LA that maintains the probability of granting permission to transmit to each of the mobile nodes under its coverage. The BS polls the mobile nodes according to these estimates and after each poll, the network feedback information is used in order to update the choice probability of each mobile node. The network feedback conveys information both on the nature of the offered traffic and the condition of the wireless link between the BS and the mobiles. Since the offered traffic is of bursty nature, when the BS realizes that the selected node had a packet to transmit, it is probable that the selected node will also have packets to transmit in the near future. Thus, its choice probability is increased. On the other hand, if the selected node notifies that it does not have buffered packets, its choice probability is reduced, since it is likely to remain in this state in the near future. The same idea applies to the condition of the wireless links. Thus, when the BS fails to receive feedback about the selected mobile's state, the mobile is probably experiencing a relatively high BER link to the BS. Since in wireless communications errors appear in bursts, the link is likely to remain in this state for the near future. Thus, the choice probability of the selected node is lowered in order to reduce the chance of a futile poll. This polling protocol, named Learning Automata based Polling (LEAP), is compared to the Randomly Addressed Polling (RAP) and Group Randomly Addressed Polling (GRAP) polling protocols for WLANs via simulation and is shown to exhibit superior performance under bursty offered traffic. Part of these results are regenerated in Figure 4, which reveals the performance superiority of LEAP in terms of throughput; a result that is attributed to its capability of adapting to the bursty production of the network offered load.

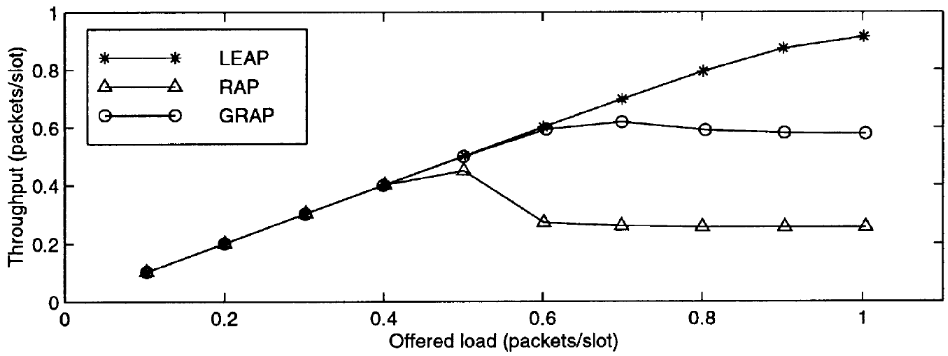

Figure 4. Throughput of LEAP versus the aggregate normalized network offered load.

In [6], the same approach is taken for an ad-hoc WLAN, thus in this case every mobile node is equipped with a $P_{LR\text{-}P}$ LA. The proposed protocol, named Ad hoc Learning-Automata-based Protocol (AHLAP), is compared to IEEE 802.11 DCF (Distributed Coordination Function) under bursty traffic conditions via simulation. Part of the simulation results are shown in Figure 5 for low and high grade burstiness of the offered traffic. One can easily see that the burstier the network traffic, the better the behaviour of AHLAP against IEEE 802.11 DCF; a fact that is attributed to the ability of AHLAP to pinpoint the active nodes in the network and grant them permission to transmit.

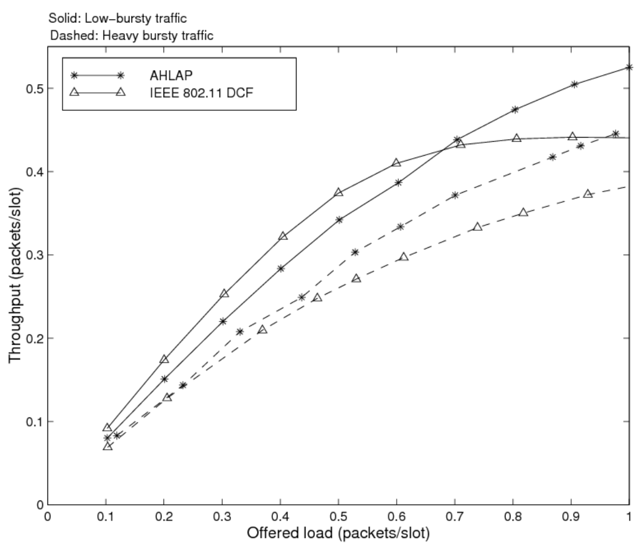

Figure 5. Throughput of AHLAP and IEEE 802.11 DCF, for low and high grade bursty traffic, versus the aggregate normalized offered load to the network.

The variability of bursty offered traffic is addressed from another perspective in [7], where the authors propose a MAC protocol for clustered wireless ad hoc networks. Inside each cluster, medium access is controlled by a designated clusterhead node via a Time Division Multiple Access (TDMA) scheme and as a result intra-cluster transmissions are collision-free. Inter-cluster communications are served by a Code

Division Multiple Access (CDMA) scheme that is overlaid on the TDMA technique, so as to achieve interference free communication between nodes in different clusters. In order for the above scheme to work, the network needs cluster formation, code assignment, and slot assignment algorithms; all three of which are based on LR-P LA. For cluster formation, each node operates a LR-P LA, which results in the organization of the nodes in a way that produces a minimum number of non-overlapping clusters. The code assignment algorithm is implemented using a LR-P LA at each clusterhead to achieve spatial reuse of the limited number of CDMA codes that can enable concurrent intercluster communication. Finally, for slot assignment to nodes, a LA at each clusterhead is used to assign to each node of the cluster a fraction of the TDMA frame proportional to the traffic load of the node. Simulation results in [7] show that the performance of proposed CDMA/TDMA scheme outperforms those of existing methods for under bursty traffic conditions.

Network layer

Learning automata have also been used in multicast routing in wireless ad-hoc networks [8]. Based on predictions of node mobility the work in [8] finds the routes with the higher lifetimes, based on which, a stochastic graph representing the virtual multicast backbone of the ad-hoc network is built. Then a distributed LA-based algorithm is applied to this graph to solve the multicast routing problem. Simulation results in [8] reveal superiority of the performance of the proposed multicast routing algorithm over existing algorithms.

Multicast routing is also addressed in [9], whose contribution is twofold: Firstly, it proposes three LA-based algorithms that find the optimal solution to the minimum weighted Steiner connected dominating set problem; secondly, one of these algorithms is implemented in a distributed manner in the network nodes to solve the multicast routing problem. Simulation results reveal the superiority of the proposed approach against well-known multicast routing protocols under a variety of performance metrics.

The virtual backbone graph mentioned in the context of [8] above is also exploited in [10] to solve the broadcast storm problem that is inherent in broadcast-based routing via global flooding. A set of LA operating on the network nodes is used to route traffic over the virtual backbone via broadcast routing. This results in a mitigation of the broadcast storm problem as the number of nodes implementing the broadcasts matches the number of the nodes in the backbone. Simulation results in [10] reveal that the proposed algorithm enjoys a significantly higher performance than that of similar existing algorithms with only a minor increase in the overhead of exchanged control messages.

Transport layer

The operation of the Transmission Control Protocol (TCP) over the error-prone wireless links can lead to significant degradation of its performance, due to the fact that TCP cannot differentiate congestion losses from wireless losses. Since the latter are the primary sources of losses in wireless networks, TCP will respond to such losses by unnecessarily invoking its congestion control mechanisms, thus degrading network performance even more. To combat this, learning automata are used in [11] in order to recognize the two kinds of losses by observing the arrival of acknowledgment and duplicate acknowledgment packets. Thus, the proposed protocol will dynamically adapt to the changing network conditions and appropriately update the *cwnd* size. Simulation results in [11] under varying wireless network conditions, show up to 18% throughput increase and 55% packet loss reduction compared to the corresponding figures of traditional TCP.

Another transport layer issue, congestion avoidance, has recently been treated with LA in the context of sensor networks tailored for healthcare applications [12]. In this approach at each intermediate node in the path from the source to the destination node, a LA continuously interacts with the environment. Based on the offered traffic at each node, each such LA adaptively learns the optimal data rate flowing through the respective node so that congestion is locally controlled at each node. Simulation results in [12] show that the proposed algorithm can efficiently avoid congestion in typical healthcare wireless sensor networks.

3.2 Applications to areas of wireless networks

Applications to ad-hoc networks

In [13] a LA-based approach is used to combat the problem of performance degradation that occurs as the size of the network grows. After proposing a LA-based centralized algorithm for finding the near optimal solution for the minimum connected dominating set representation of the network, the authors propose a distributed version of the centralized algorithm that can be implemented at every network node. Simulation results reveal the superiority of the proposed approach over the existing ones.

Applications to sensor networks

The primary concern of wireless sensor networks is to reduce energy consumption. This can be achieved by data aggregation, which is performed in intermediate nodes en-route to the sink. Data aggregation prevents redundant information being forwarded by neighboring nodes that sense the same event. The benefits of aggregation are maximized when the network routes data packets over routes comprising nodes with a high probability of having sensed the same information with that carried in the routed packet. Since these paths are not static, but rather follow changes in the environment, [14] proposes an aggregation scheme where each node is equipped with a LA in order to collectively learn the path of aggregation with maximum aggregation ratio. Simulation results in [14] show that the proposed method exhibits significantly better performance against other aggregation mechanisms, especially in dynamic environments.

LA have also been used in [15]. This work proposes a fully distributed LA-based algorithm that efficiently clusters wireless sensor networks in order to reduce energy consumption and increase network lifetime. Contrary to other approaches which recluster the network at regular time intervals, in the proposed algorithm reclustering is performed locally when network conditions mandate it, thus resulting to a decrease in energy consumption. Simulation results in [15] reveal increased network lifetime and energy efficiency when compared to other clustering methods.

In [16] the authors propose LA-based a scheduling algorithm to deal with the problem of using the fewest possible number of sensors to detect moving objects via a sensor network. This objective of course translates to reduced energy consumption for the network. According to the proposed approach, each sensor node is equipped with a set of LA. The objective of the each such set is to learn the maximum sleep duration for the node so that the detection rate of targets by the node does not fall dramatically. Simulation results in [16] reveal increased energy efficiency over existing methods.

Cognitive radio

LA can also provide for a powerful learning tool in the environment of cognitive wireless networks. Reference [17] studies the problem of Quality of Service (QoS) enabled opportunistic spectrum access for non-licensed users. For the case when not all licensed users can be supported over the available spectrum, each node should operate a LA. A set of priorities is defined and the probability that a node can access the medium is set to be proportional to the node’s priority. Then, a LA-based algorithm is used to solve the priority-based medium access in a distributed way.

In [18], the authors proposes a method of managing the network routes that carry traffic between non-licensed users in a cognitive multihop network. The problem addressed is attributed to the randomness that derives from an unpredictable behaviour by the licensed users. In the proposed approach, LA are used at the network nodes in order to find those routes that are the least affected by licensed users.

Wireless data broadcasting

In the area of wireless data broadcasting, [19] proposed a LA-based push system that enables adaptivity in environments with dynamic and a priori unknown client demands. In such environments, there exists a large number of clients with overlapping demands for the various items that reside in the server's database, which are delivered to the clients via broadcasting. In the system presented in [19], a S_{LR-I} LA is incorporated at the broadcast server which uses simple feedback from the clients to obtain an estimate of the client demands. Simulation results in [19] show that under dynamic environments, where client demands change over time show that the adaptive system provides superior performance when compared to the non-adaptive approach. Part of these results are depicted in Figure 6 for a server broadcasting 300 items and show that the performance gains in terms of mean waiting time for an item (response time). These gains for the adaptive system increase for increasing values of the data skew coefficient; a parameter that reflects the amount of commonality in the demands of the various clients.

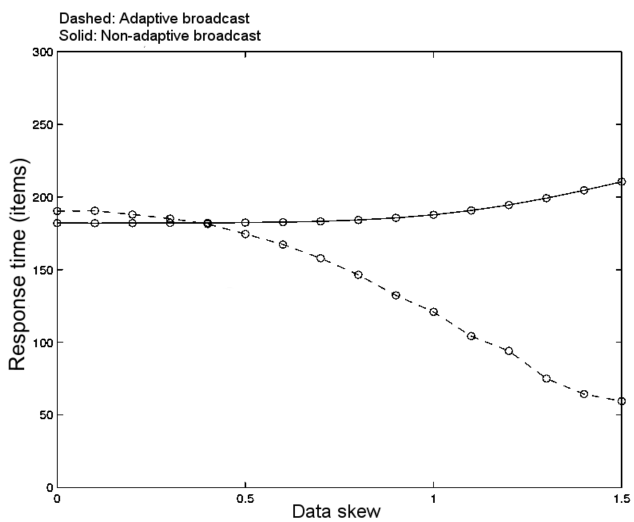

Figure 6: Performance gains of the method of [19] due to the incorporation of adaptivity.

4. Conclusions

Adaptivity to the various unknown and time varying aspects of the wireless environment is a challenging research topic and has arised as one of the most desired aspects of protocols for wireless networks. Learning Automata are Artificial Intelligence tools that have been used towards achieving the goal of such adaptivity. After discussing the dynamic properties of the wireless environment that mostly affect network performance, this paper overviews the structure, operation and key parameters of LA as well as the impact of these parameters in their behavior. Next, it provides an up-to-date survey of the state-of-the art approaches in using Learning Automata to provide adaptivity to several layers of the wireless networking stack, namely physical, MAC, routing and transport layers and additionally on specific wireless networking environments, such as sensor networks and wireless data broadcasting. The goal of this survey is to summarize research in the area and bring forward the potential of LA as an important tool in the research towards intelligent wireless networking.

References

[1] M.S. Obaidat**,** G.I. Papadimitriou and A.S. Pomportsis, "Learning Automata: Theory, Paradigms and Applications", *IEEE Transactions on Systems, Man and Cybernetics-Part B*, vol.32, no.6, December 2002, pp. 706-709.

[2] T. Joshi, D. Ahuja, D.Singh and D. Pagrawal, "SARA: Stochastic Automata Rate Adaptation for IEEE 802.11 Networks", *IEEE Transactions on Parallel and Distributed Systems*, vol.19, no.10, October 2008, pp.1579–1590.

[3] M.A. Haleem, R. Chandramoulli, "Adaptive Downlink Scheduling and Rate Selection: a Cross-layer Design", *IEEE Journal On Selected Areas in Communications*, vol.23, no.6, June 2005, pp.1297-1297

[4] Y. Xing, R. Chandramouli, "Stochastic Learning Solution for Distributed Discrete Power Control Game in Wireless Data Networks", *IEEE/ACM Transactions on Networking*, vol.16, no.4, August 2008, pp.932-944.

[5] P. Nicopolitidis, G.I. Papadimitriou and A.S. Pomportsis, "Learning-Automata-Based Polling Protocols for Wireless LANs", *IEEE Transactions on Communications*, vol.51, no.3, March 2003, pp.453-463.

[6] P. Nicopolitidis, G.I. Papadimitriou and A.S. Pomportsis, "Distributed Protocols for Ad-Hoc Wireless LANs: A Learning-Automata-Based Approach", *Ad Hoc Networks*, Elsevier, vol.2, no.4, October 2004, pp.419-431.

[7] J. Akbari Torkestani and M.R. Meybodi, "An Efficient Cluster-based CDMA/TDMA Scheme for Wireless Mobile Ad-hoc Networks: A Learning Automata Approach", *Journal of Network and Computer Applications*, Elsevier, vol.33, no.4, July 2010, pp.477-490.

[8] J. Akbari Torkestani and M.R. Meybodi, “Mobility-based Multicast Routing Algorithm for Wireless Mobile Ad-hoc Networks: A Learning Automata Approach”, *Computer Communications*, Elsevier, vol.33, no.6, April 2010, pp.721-735.

[9] J. Akbari Torkestani and M.R. Meybodi, “Weighted Steiner Connected Dominating Set and its Application to Multicast Routing in Wireless MANETs”, *Wireless Personal Communications*, Springer, February 2010.

[10] J. Akbari Torkestani and M.R. Meybodi, “An Intelligent Backbone Formation Algorithm for Wireless Ad hoc Networks based on Distributed Learning Automata”, *Computer Networks,* Elsevier, vol.54, no.5, April 2010, pp.826-843.

[11] B. Venkata Ramana and C. Siva Ram Murthy, “Learning-TCP: A Novel Learning Automata Based Congestion Window Updating Mechanism for Ad hoc Wireless Networks”, *Lecture Notes in Computer Science,* Springer, vol.3769, 2005, pp.454-464.

[12] S. Misra, V. Tiwari, M.S. Obaidat, “LACAS: Learning Automata-based Congestion Avoidance Scheme for Healthcare Wireless Sensor Networks”, *IEEE Journal on Selected Areas in Communications*, vol.27, no.4, May 2009, pp.466-479.

[13] J. Akbari Torkestani and M.R. Meybodi, “Clustering the Wireless Ad-Hoc Networks: A Distributed Learning Automata Approach”, *Journal of Parallel and Distributed Computing*, Elsevier, vol.70, no.4, April 2010, pp.394-405.

[14] M. Esnaashari and M.R. Meybodi, “Data Aggregation in Sensor Networks using Learning Automata”, *Wireless Networks*, Springer, vol.16, no.3, April 2010, pp.687-699.

[15] M. Esnaashari and M.R. Meybodi, “A Cellular Learning Automata based Clustering Algorithm for Wireless Sensor Networks”, *Sensor Letters*, Vol. 6, No. 5, October 2008, pp. 723-735.

[16] M.Esnaashari and M.R. Meybodi, "A Learning Automata based Scheduling Solution to the Dynamic Point Coverage Problem in Wireless Sensor Networks", *Computer Networks*, Elsevier, 2010, to appear.

[17] Y. Xing, C.N. Mathur, M.A. Haleem, R. Chandramouli, K.P. Subbalakshmi, “Dynamic Spectrum Access with QoS and Interference Temperature Constraints”, *IEEE Transactions on Mobile Computing*, vol.6, no.4, April 2007, pp.423-433.

[18] Y. Song, C. Zhang and Y. Fang, “Stochastic Traffic Engineering in Multihop Cognitive Wireless Mesh Networks”, *IEEE Transactions on Mobile Computing*, vol.9, no.3, March, 2009, pp.305-316.

[19] P. Nicopolitidis, G. I. Papadimitriou and A.S. Pomportsis, “Using Learning Automata for Adaptive Push-Based Data Broadcasting in Asymmetric Wireless Environments”, *IEEE Transactions on Vehicular Technology*, vol.51, no.6, , November 2002, pp.1652-1660.