Dynamic Network State Learning Model for Mobility Based WMSN Routing Protocol

A.Ajina[1], Mydhili K.Nair[2]

[1]Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology, Bangalore, India.
[2]Department of Information Science and Engineering, Ramaiah Institute of Technology, Bangalore, India.

***Abstract*:** The rising demand of wireless multimedia sensor networks (WMSNs) has motivated academia-industries to develop energy efficient, Quality of Service (QoS) and delay sensitive communication systems to meet major real-world demands like multimedia broadcast, security and surveillance systems, intelligent transport system, etc. Typically, energy efficiency, QoS and delay sensitive transmission are the inevitable requirements of WMSNs. Majority of the existing approaches either use physical layer or system level schemes that individually can't assure optimal transmission decision to meet the demand. The cumulative efficiency of physical layer power control, adaptive modulation and coding and system level dynamic power management (DPM) are found significant to achieve these demands. With this motivation, in this paper a unified model is derived using enhanced reinforcement learning and stochastic optimization method. Exploiting physical as well as system level network state information, our proposed dynamic network state learning model (NSLM) applies stochastic optimization to learn network state-activity that derives an optimal DPM policy and PHY switching scheduling. NSLM applies known as well as unknown network state variables to derive transmission and PHY switching policy, where it considers DPM as constrained Markov decision process (MDP) problem. Here, the use of Hidden Markov Model and Lagrangian relaxation has made NSLM convergence swift that assures delay-sensitive, QoS enriched, and bandwidth and energy efficient transmission for WMSN under uncertain network conditions. Our proposed NSLM DPM model has outperformed traditional Q-Learning based DPM in terms of buffer cost, holding cost, overflow, energy consumption and bandwidth utilization.

***Keywords*:** Mobility, Wireless Multimedia Sensor Network, Network state learning, Dynamic power management, Q-learning, Markov decision process**.**

1. Introduction

In recent years, the high pace rises in wireless communication systems and associated application demands have motivated researchers to achieve low cost, efficient and robust routing protocols. The ease of implementation, efficiency and scalability of wireless communication systems makes it the dominating communication system solution. Among varied wireless technologies, Wireless Sensor Network (WSN) is one of the most used technologies to serve an array of applications including surveillance facilities, industrial monitoring and control, wireless broadcast, defense purposes, traffic surveillance etc. Being a low-cost solution for an array of communication purposes, WSN has gained widespread interest and recently has given rise to a new domain called Wireless Multimedia Sensor Network (WMSN) that enables multimedia data transmission over it. WMSN has significantly high serving capacity for multimedia real time data (RTD) transmission for multimedia broadcast and decision purposes. WMSN comprises multiple sensor nodes interfaced through wireless communication media that collect event driven real time multimedia data from transmitter and forwards it to the sink for decision process [1]. Interestingly, WMSNs have been found significant for emerging technologies such as Internet of Things (IoTs) [2-4], machine learning where it requires enabling energy efficiency, minimal end-to-end delay, bandwidth utilization and QoS delivery. In practice most of WMSN systems undergo adversaries like computational overheads, energy consumption, network contention, congestion and data drop, etc. Furthermore, the probability of computational overhead, buffer cost or the holding period etc., can't be ignored. Therefore, it demands certain efficient protocol to enrich WMSN to avoid mentioned adversaries [5,6]. Though, numerous efforts have been made, most of the existing routing approaches are found limited to provide timely data delivery, delay sensitive RTD delivery, energy and bandwidth efficient communication [7,8]. Multimedia communication over WMSN often demands minimal latency, delay sensitive transmission, higher throughput, higher bandwidth utilization, and low energy consumption.

Recently, mobility-based sensor networks have gained significant attention due to its data gathering efficiency, energy reduction ability, and higher scalability. Due to excessive mobility there is often uncertainty in the network that makes forwarding decision intricate that as a result makes QoS delivery difficult. Median Access Control (MAC) optimization in physical layer transmission scheduling can enable better performance. However, such single layer optimization doesn't yield optimal solution and requires certain cross layer architecture-based routing [8-11] to maintain overall functional efficacy of the network. Authors [12] developed an interference aware routing protocol for WMSN. Cross-layer architecture was developed by incorporating an integrated model connecting all layers of the protocol stack (IEEE 802.15.4) that was found efficient to achieve QoS [13]. Still, the prime limitation is the use of single network parameter for forwarding decision. Unfortunately, as per our present knowledge not much significant efforts are made to incorporate mobility with WMSN. The prime reasons of such limitations are topological dynamism that cumulatively makes transmission decision intricate. To deal with it, we developed a multi-constraints cross-layer model named network condition aware routing protocol for WMSN (NCAM-RP) [14]. NCAM-RP focused on integrating mobility with WMSN while ensuring QoS delivery, higher throughput, deadline sensitive transmission and low data drop. In addition, it

emphasized on achieving higher RTD delivery, while maintaining best feasible resource provision for Non-real time (NRT) data.
Mobility in WMSN introduces topological variations in the network conditions that as a result forces network to undergo buffer contingency, contention, data drop, delay and energy exhaustion [12,14]. On the contrary, multimedia communication often requires timely data delivery with minimum resource consumption [5,6]. Undeniably, transmission over WMSN requires minimal buffer or holding period (the time during which a node retains or holds packet before transmission and delivery to the sink), higher resource utilization, and minimal energy consumption at each node. Practically, enabling delay resilient and energy efficient transmission are directly related to power control at PHY. Literatures reveal that DPM at system level and PHY switching based power control at physical layer can cumulatively enable swift data transmission, optimal bandwidth utilization, and more importantly energy-efficient adaptive transmission [15,16]. However, in dynamic network topology performing transmission scheduling and PHY switching is highly intricate and even infeasible unless a controller is aware of network conditions and channel states. In this case, performing network state learning is always a dominating solution [16]. In mobile WMSN network parameters can vary over time and hence learning the known as well as unknown parameters (Ex. Buffer available at a node, deadline time etc.,) and adapting PHY switching accordingly can make overall routing efficient. Since, the network parameters can be in large amount and even varying over time, performing fast learning or convergence can make overall network time and QoS efficient. The classical learning methods such as Q-learning [15,17,18] often uses known parameters to make any stochastic decision. On contrary, mobile-WMSN requires a suitable learning approach that could exploit or apply both known as well as unknown parameters to make optimal PHY switching decision to achieve timely, resource efficient and QoS communication. Considering it as motivation, in this paper a robust network condition or network state learning based stochastic model is developed for DPM and PHY switching. Majority of existing approaches have applied either system level DPM or physical layer control such as Adaptive Modulation and Coding (AMC) [16]. However, these techniques as individual solution can't assure optimal solution. Unlike traditional approaches [16-19], in our proposed method we have exploited the efficiency of system level DPM and physical layer control AMC. Thus, the proposed model can be stated as a cross-layer approach comprising both the system level design as well as physical layer power control. It can make the proposed approach suitable for mobile-WMSN. In our proposed model, we have developed a novel stochastic learning approach that learns over known as well as unknown network parameters to make adaptive DPM and PHY switching to achieve time, resource and quality efficient communication over sensor network. Considering the above features, here onwards we state our proposed model as Network State Learning Machine (NSLM) based DPM. The predictive characteristic of NSLM enables it to be called as a stochastic model that intends to achieve optimal DPM and transmission scheduling. To behave as stochastic model, NSLM exploits Markov Decision Process (MDP) using HMM and Lagrange relaxation that assures swift convergence for timely transmission decision. Thus, our proposed system can be called as the constrained MDP problem, where emphasis is made on achieving swift learning and PHY switching to have timely and resource efficient transmission. The proposed system is developed using MATLAB tool, where simulation is made to assess effectiveness in terms of delay, power, resource consumption, convergence rate etc. To examine performance of the proposed system, we have compared it with traditional Q-Learning based DPM, where the overall results affirm that our proposed NSLM technique outperforms classical Q-Learning based power management and transmission scheduling for WMSN. The results and the cumulative robustness of our proposed technique enable it to be suitable for mobile WMSNs.
The other sections of the presented paper are divided as follows: Section 2 presents the related work; Section 3 discusses the proposed research work, which is followed by results discussion in Section 4. Conclusion and future scopes are presented in Section 5. References used in this paper are given at the end of the manuscript.

2. Related Work

To derive a novel WMSN solution, enriching a protocol with energy efficiency, minimal delay and QoS delivery enriched communication in sensor network is must. To achieve it authors [12,13] derived a multi-relay-based delay-sensitive transmission for wireless communication. They [12] found that their approach is effective in reducing energy exhaustion. Considering significance of PHY layer for energy enhancement, authors [20, 21] examined LEACH and AODV routing protocols. Authors [22] applied CSMA/MAC based iQueue-MAC to deal with dense traffic conditions and transmission scheduling to achieve energy efficiency. The effort to use local network statistics such as node distance, successful transmission probability, channel statistics etc., were used to develop energy efficient routing protocol [23]. In [24] cross-layer architecture-based packet-forwarding scheme named channel-aware geographic-informed forwarding (CAGIF) was developed; however, the issue of topological variations and network uncertainty could not be addressed. For data sensitive transmission, authors [24] proposed "green Task-Based Sensing" (gTBS) where they focused on sleep-awake scheduling to reduce energy consumption. Authors [25] proposed a cross layer model where they considered application layer, network, MAC and PHY layer to achieve energy efficiency. Authors applied topological states including node's current location to perform route discovery. They applied MAC to exhibit power control for reduction in energy consumption. In [26], network layer and adaptive power control approaches were suggested to reduce energy consumption. To deal with multiple constraints a convex optimization model was developed in [27], which was later used to design a cross-layer model to increase network lifetime and bandwidth. In [28] PHY, MAC and network layers were applied to derive a cross layer formulation, where authors applied residual energy and link quality as the network parameters and applying Fuzzy learning they determined next hop decision to achieve low energy consumption. Considering the need of a network learning approach for best forwarding path estimation, authors [29] developed a hybrid evolutionary approach using genetic algorithm (GA) and bacteria foraging optimization (BFO). However, could not address the issue of dynamic or uncertain network condition. Interestingly those approaches are confined to provide optimal solution with uncertain network conditions. In [30], a mobile agent-based network learning model was developed for energy efficient routing.

Authors made effort to derive a cross layer model-based routing protocol for bandwidth constrained [1] and QoS demanding [31] WMSNs. To achieve network condition adaptive data transmission, researchers [32] derived an entropy tracking model which uses network condition variations and the correlation structure (based on distance between nodes, characterizing spatial correlation between multimedia sensors) to define its multi-rate transmission scheduling. Authors focused on PHY scheduling in a way that the transmission rate could be adjusted based on deadline time and data rate required. In [19] a cross-layer model was developed; however, they could not address the issue of data priority and adaptive transmission. Adaptive link quality based WMSN routing model were developed in [33] and [34]. Researchers [30, 35] found geographic routing protocols efficient for WMSNs due to the shortest path-based transmission scheduling. However, it doesn't guarantee performance with varying topology and uncertain network conditions, particularly during congestion or contention scenarios. A multipath routing approach was suggested in [8] to achieve delay sensitive transmission and per node holding time reduction. In [36], authors developed a cross-layer model-based MAC optimization scheme where they focused on RTD delivery by means of multiple paths. Authors [37] made effort to deliver QoS in WMSN, however at the cost of the energy consumption. In [4, 11], authors derived a QoS aware MAC routing protocol where they applied bandwidth and delay as the WMSN forwarding node selection criteria. In multimedia transmission, assigning data priority is must and hence with this objective, authors [38] derived a doubling-based service differentiation that mainly focused on prioritized data transmission over sensor network, though the QOS constraints and energy efficiency remained unexplored. On contrary, authors [39] derived RTD delivery model under resource constrained scenarios. Deadline time-based traffic prioritization was done in [40]; however, could not address the fundamental WMSN requires such as bandwidth utilization and energy exhaustion or node life span. Authors [41] developed a multipath routing scheme where they used traffic priority [42], link quality and energy to perform MAC scheduling. Considering efficient power management schemes for wireless sensor networks, most of approaches have either used power control and adaptive modulation and coding (AMC) at physical layer or dynamic power management at system layer. However, along these approaches are confined and hence requires a unified model to enable optimal DPM and PHY switching system to yield delay sensitive, energy and bandwidth efficient routing protocol for WMSNs.

Considering the significance of a robust routing protocol for wireless multimedia sensor network (WMSN), in our previous work [14], we developed a novel multi-constraint, cross-layered routing protocol. With intend to develop a low cost and efficient WMSN solution we developed mobility based WMSN protocol NCAM-RP. It was implemented at Application layer, MAC layer and Network layer of the WMSN protocol stack. NCAM-RP protocol was emphasized on delivering higher throughput for mission critical RTD data transmission, while ensuring fair resource provision to the NRT data over WMSNs. Though, it exhibited better in terms of throughput and delay sensitive transmission but the key requirement such as bandwidth efficiency, energy efficiency and delay could not be addressed. Strengthening NCAM-RP with DPM and network adaptive PHY switching can meet QoS and energy efficient demand. NCAM-RP being a mobile-WMSN approach requires dealing with the uncertain network conditions caused due to dynamic (i.e., mobile) topology, therefore it requires certain stochastic (network) learning model for DPM and power control policy. Considering it as motivation, in this paper we have developed a robust dynamic network state learning model (NSLM). It can be considered as an enhanced reinforcement learning where unlike generic Q-Learning based approach, we have considered both the known as well as unknown network parameters to make transmission switching decision and DPM policy. The overall proposed DPM function is considered as constrained Markov Decision Process (MDP), where Hidden Markov Model (HMM) and Lagrangian relaxation have been applied to enhance convergence rate that enables swift decision policy estimation for delay-sensitive DPM and transmission scheduling. Our proposed model encompasses both the stochastic optimization as well as enhanced reinforcement learning based online network state learning to derive optimal transmission or DPM policy. To enable a cross layer solution, our proposed NSLM model incorporates both physical layer components such as power control and AMC as well as system level components such as DPM. It enables our model to exploit the physical as well as system level information to derive a novel DPM solution or transmission solution. In proposed unified stochastic optimization and learning model based DPM, the time series (network) events have been learnt, based on this an optimal state-action policy is derived for DPM and PHY switching or transmission scheduling. The discussion of the overall proposed model and its implementation is given in the next section.

3. System Model

This section discusses the proposed NSLM based unified DPM and transmission scheduling (PHY switching). As the proposed approach considers network condition statistics to control transmission switching, it employs both the PHY layer as well as system layers components. Considering network condition learning, NSLM considers time-slotted model where the overall simulation time is divided into a defined time interval (say, (Δt)). Here, the time division is decided in way that certain n^{th} time-slot is defined in the form of $(n\Delta t, (n+1)\,\Delta t\,)$. NSLM performs transmission decisions at the start of each time-slot. Here, the network state has been obtained at the interval of Δt. The channel states variables or network state variables (NSVs) such as buffer availability, bit error probability (BEP), etc are estimated for each slot Δt. One of the key novelties of the proposed work is the consideration of the known as well as unknown state variables to learn network conditions and derive optimal PHY scheduling and transmission decision policy. Thus, the overall proposed model (Figure 1) intends to schedule PHY switching and transmission scheduling to accomplish optimal bandwidth utilization, low buffer holding period, minimal delay and energy consumption.

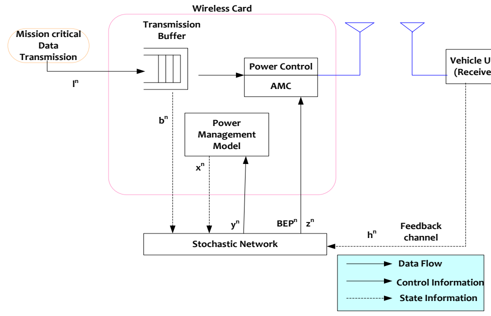

Figure 1. Proposed network condition adpative DPM and PHY scheduling model

The following sub-sections discuss the proposed network condition aware DPM model.

3.1 PHY Layer DPM and transmission Scheduling

The present day physical layer centric approaches apply AMC based power control model to enable optimal power management and (PHY switching) transmission scheduling. Most of the multimedia communication systems apply AMC based power management policy. It signifies the robustness of the AMC strategies for QoS achievement. This approach plays vital in reducing power consumption, however retaining system as active even during the phase when there is no transmission, costs a lot and thus reduces overall efficiency. The use of system level components such as wireless card can assist scheduler to known real time network state to make DPM. We exploit this approach to derive network condition aware DPM model for WMSN. The proposed model considers the frequency non-selecting channels where $c^n \in \mathbb{C}$ refers the channel-fading-coefficient (CFC) during time slot n. Here, the channel states $\mathbb{C}$ can be considered as discrete and finite variable [43-46]. For computation, $\mathbb{C}$ is constant during a fixed interval n that makes channel state $\{c^n \in \mathbb{C} : n = 0,1,\dots\}$, and allied computation easy. Literatures suggest that the varying network states such as dynamic channel conditions, buffer availability, channel state information (CSI), BEP etc. can be obtained by means of Markov model with the transmission probability of $t_p{}^h(c' | c)$. Here, physical layer is considered as single carrier SISO (Single Input Single Output) with fixed data rate. The rate of data generation is presented $(1 / \mathbb{T}_s)$ that also signifies symbols per generation. The transmitter node is supposed to be generating data at the rate of $a^n \geq \mathbb{T}_s$(bits/s), where $a^n \geq 1$ presents the overall bits obtained by AMC's modulation. We consider each packet of $\mathcal{N}$ bits, where $\mathbb{T}_s$ is known and fixed. To measure the BEP (α), receivers use detectors (maximum likelihood α^n) (1) with the transmission power E_{tx}^n (2). Mathematically,

$$\alpha^n = \alpha\,(c^n, E_{tx}^n, th^n)\ a\,v\delta \tag{1}$$

$$E_{tx}^n = E_{tx}\,(c^n, \alpha^n, th^n)\ , \tag{2}$$

where th^n refers the throughput for each slot n. Considering independent bit errors across the WMSN network, the loss rate for the individual packet of $\mathcal{N}$ bit is measured in terms of BEP. Thus, the packet loss rate (PLR) is estimated using (3).

$$\mathrm{PLR}^n = 1 - (1 - \alpha^n)^{\mathcal{N}} \tag{3}$$

In WMSN, the throughput being one of the most vital parameter for QoS delivery, characterizes the dynamic link quality between sensor nodes. Considering throughput as the significant network parameter, we have considered it as the decision variable. Consider a case to transmit $\mathcal{N}th^n$ for a time slot Δt (to achieve a cumulative throughput of $th^n \in \mathbb{TH}$), it is must to recognize the total bits per symbol (α^n)during modulation. Here, α^n is obtained as (4).

$$\alpha^n = \lceil th^n \mathcal{N} \mathbb{T}_s / \Delta t \rceil\ , \tag{4}$$

The known packet throughput with α^n, and CSI (c^n), the decision towards transmission power $(\sum_{tx}^n)$ becomes easy which can be further applied to estimate BEP or vice versa. In QoS oriented WMSN, BEP can be a suitable decision parameter and can be applied for transmission power decision which is anticipated for DPM and PHY switching. PHY switching and power transmission scheduling can be feasible with known BEP and E_{tx}^n,; however, with mobility WMSN the varying topology could cause these parameter to vary over time and hence remains unknown. This condition is strongly in conjunction with NCAM-RP routing protocol [14]. To deal with this situation, we intend to derive a dynamic network state learning model to enable adaptive PHY switching or DPM to adapt AMC model. Noticeably, implementing simultaneous modulation and coding mechanism the equations (1-2) can force outputs (1-2) unachievable. To avoid it and to make the equations (1) and (2) known, fitting the network parameters can be vital [47-49]. NSLM applied an enhanced reinforcement learning approach based stochastic optimization and learning model to derive optimal state-activity policy for enhanced DPM. NSLM DPM considers both system level as well as PHY level network parameters. System layer parameters typically considers higher layer of the protocol stack and is armored with wireless card to provide dynamic channel states or information to make proper decision.

3.2 System-Level Model for DPM

In addition to the DPM and AMC, which measure the active transmission power, we intend to introduce certain additional wireless components (Figure. 1) in the low power states that consequently can reduce power consumption. In general, the wireless cards of the major sophisticated communication system resides in the power state set $p = \{\mathrm{ON}, \mathrm{OFF}\}$ which can be scheduled for sleep and awake states. The sleep and awake state can be well understood in terms of ON and OFF states respectively. Let the switching functions be $s = \{S\_ON, S\_OFF\}$. A channel with respective state conditions c characterizes the state p of the wireless card. Let α^n be the probability of the maximum BEP, s be the power control action to achieve dynamic power control, DPM or PHY switching. The variable th signifies the overall throughput. Thus, using these variables, the power required for transmission and link adaptive transmission scheduling can be obtained by (5).

$$\tau([c,p],\alpha,s,th) = \begin{cases} [E_{on}+E_{tx}(c^n,\alpha^n,th^n)], & if\ p=on, s=s\_on \\ E_{off}, & if\ p=off, s=s\_off \\ E_{tr}, & otherwise \end{cases} \quad (5)$$

Here, E_{tx} is the transmission power in Watts unit. The other variables E_{on} and E_{off} signify the power consumed during state conditions, "ON" and "OFF", respectively. The variable E_{tr} presents the energy consumed during the transition (i.e., from ON to OFF or vice versa). Here, we assume that $E_{tr} \geq E_{on} \geq E_{off} \geq 0$ such that there can be certain penalty to perform switching in between the two states [50]. Researchers [50] suggested that the power management states sequence $\{p^n \in \mathbb{P} : n = 0,1,\ldots\}$ can be derived in the form of a controlled Markov model having the distinct transition likelihood $m^p(p'|p,s)$. Let $\mathcal{M}^p(s)$ be the transition likelihood matrix (TLM) conditioned on y in the way that $\mathcal{M}^p(s) = [m^p(p'|p,s)]_{p,p'}$. Due to the high level of abstraction, there could be the likelihood of a non-deterministic delay in combination with the DPM state transition, and thus

$$\mathcal{M}^p(s\_on) = \begin{array}{c} \\ on \\ off \end{array} \begin{pmatrix} on & off \\ 1 & 0 \\ \theta & 1-\theta \end{pmatrix}$$
$$\mathcal{M}^p(s\_off) = \begin{array}{c} \\ on \\ off \end{array} \begin{pmatrix} on & off \\ 1-\theta & \theta \\ 0 & 1 \end{pmatrix} \quad (6)$$

In (6), the state conditions presented in row and column refers the present (p^n) and the next state (p^{n+1}), respectively. The variable $\theta \epsilon (0,1)(\text{resp.}\ 1-\theta)$ refers the successful transmission probability. In our model, to maintain elucidated solution, we have considered the power state transition (PST) as deterministic $(\theta = 1)$ and is applied with condition $\theta < 1$ (where θ can be known or unknown). NSLM assures non-zero throughput th, when $p =$ ON and $s =$ S_ON, otherwise throughput th is obtained as zero.

3.3 Transmission Model and Buffer Modelling

Considering functional approach, similar to our previous work [14] where RTD was stored in prioritized buffer while NRT data were queued in first-in first-out (FIFO) manner, for simple implementation we consider FIFO transmission buffer (Figure. 1). Here, the transmitter transmits η^n packets into the associated buffer in each slot n, where η^n has the distribution $m^n(\eta)$. Each packet is of $\mathcal{M}$ bits size and the receiving process $\{\eta^n : n = 0,1,\ldots\}$ is considered being self-directed and is uniformly distributed across time slot n. The received packets are stored in a fixed size buffer, which can store the maximum of $\mathbb{Q}$ packets. The buffer state $b_s \in \mathbb{P}_{\mathbb{P}\in 0,1,2,\ldots,\mathbb{P}}$ emerges iteratively using (7).

$$q^0 = q_{init} \quad (7)$$
$$q^{n+1} = min(q^n - f^n(\alpha^n, th^n) + \eta^n, \mathbb{Q})$$

Where q_{init} signifies the initial buffer state and $f^n(\alpha^n, th^n) \leq th^n$ is the goodput during each time slot. Here, good-put refers the total number of packets transmitted successfully. Typically, good-put depends on two parameters, BEP $(\alpha^n \in \varepsilon)$ and throughput $th^n \in \{0,\ldots,\min(q^n, th_{max})\} = th(q^n)$.

To achieve delay sensitive delivery for WMSN, buffer cost is introduced that rewards the system for minimizing overall delay, particularly queuing delay. It also plays vital role in avoiding packet overflows caused due to abrupt escalation in transmission delay. It might occur frequently in mobility based WMSNs. Varying network dynamicity, fading, and traffic burst are the predominant reasons for the packet overflow conditions. Buffer cost is considered as the expected sum of the overall Holding Cost (HC) and the Overflow Cost (OC) in conjunction with the received packet and the good put distribution. Mathematically.

$$d([q,m],\alpha,s,th) = \sum_{\eta=0}^{\infty}\sum_{f=0}^{o} m^{\eta}(\eta)m^{f}(f\mid\alpha,th)\left\{\underbrace{[q-f]}_{holding\ cost} + \underbrace{n max([q-f]+\eta-\mathbb{Q},0)}_{overflow\ cost}\right\} \quad (8)$$

In (8) the Holding Cost (HC) refers the total packets which could not be delivered successful and is available in the buffer (initially). In fact, it is equivalent to the buffer cost. In case of stable buffer condition (i.e., no buffer overflow), HC is proportional to the queuing delay [51]. In contrast, for unstable buffer condition (i.e., overflow condition) the proposed overflow cost introduces a penalty factor (n) for each dropped packet. Here, n plays vital role in solving the optimization problem.

3.4 NSLM DPM as Constrained Markov Decision Problem

This section primarily discusses the implementation of the proposed NSLM model for DPM and PHY switching policy derivation.

3.4.1 Constrained HMM Design and Decision Mechanism

Our proposed model at first employs a joint state model that contains key network information such as buffer state information, power state, and link quality related information. Furthermore, we introduced a vector called Joint Action Vector (JAV) that specifically contains BEP, DPM action y, and the network throughput per slot th.. (here onwards we present throughput per slot by h). Mathematically,

$$j \triangleq (\alpha, s, th) \in \mathbb{J} \quad (9)$$

We formulate WMSN network state as a sequential set given by $\{y^n : n = 0,1,\ldots\}$, derived as a controlled MDP. In applied MDP scenario, the state transition probability is calculated based on certain conditionally independent parameters (i.e., buffer conditions or availability, DPM state transitions).

In NSLM, MDP applies decremental criterion as it needs a single optimality function which is independent of MDP chain model. Decremental objectives not only enable swift computation but also enable cost effective solution in comparison to other state of art technique such as Q-learning. In practice, the traffic conditions of the WMSN and associated channel states are stationary only for a small-time period; and therefore can predict costs which DPM policy could use to make future decision. In this manner, the expected future costs can be significantly minimized during DPM learning and decision for each time slot n. It is significant to optimize the time average cost, because of unknown lifetime of an application, where there can be unknown prior information. Considering this fact, we have assumed that the application could stop with the probability of $(1-\rho)$ in n time slot and thus the MDP would turn into zero-cost power use state. In continuation of the above discussed optimization objectives and allied constraints, a penalty factor called per-packet penalty (PPP) n is estimated that is typically caused due to overflow in buffer cost function (9). To have a better solution, the penalty n must be optimally large so as to maintain minimal packet drop during transmission, which is must for multimedia communication over WMSNs. To achieve sub-optimal packet drop, we have applied penalization for each dropped packet by minimal cost that could induce it to the buffer for further transmission. The maximum cost that a packet may achieve when entering the buffer is the infinite holding cost (HC) incurred when the packets are hold for long time or forever.

Typically, in WMSN systems, a data packet reaching a node at n_0 doesn't inevitably add any significant holding cost till (n_0+1), and hence the infinite HC can be obtained as (10).

$$n = \sum_{n=n_0+1}^{\infty} (\rho)|^{n-n_0} = \frac{\rho}{1-\rho} \tag{10}$$

Where ρ presents the reductive factor (RF) or the decremental factor (DF).

3.4.2 Lagrangian Relaxation Model

The optimization model derived based on minimize decremental power cost can be re-modeled as certain unimpeded MDP by means of Lagrange relaxation, also called Lagrange multiplier. The Lagrange multiplier can be used in relation to the delay parameter. We have derived a factor called Lagrangian cost function (LCF) (11):

$$h^{\mu}(y,j) = \tau(y,j) + \mu d(y,j) \tag{11}$$

where the value $\mu \geq 0$ signifies the Lagrange relaxation or multiplier. $\tau(y,j)$ presents the power cost as derived in (5) and $d(y,j)$signifies buffer cost (11).

The following section discusses the brief of the DPM learning policy.

The overall relationship between varying WMSN network states and policies or the actions at the time slot n, and the network state at the $n+1$ time slot are presented in Figure 2. It also depicts the network state transition probability at $(n+1)$ and associated SVF and cost functions etc.

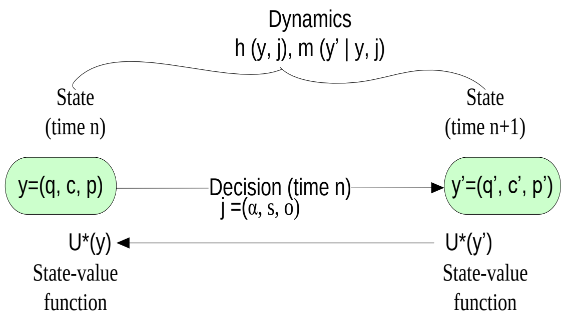

Figure 2. Visualization of the MDP

3.4.3 DPM Learning Model

In the proposed model, we have considered the buffer cost and the transition probability function (TPF) as the unknown variables, as it varies due to topological variations [14] As a result, it makes the measurement of the key network parameters $U^*$and ψ^*possible through value iteration. These network parameters are needed to be learnt during active life time (i.e., online) based on the past experiences or WMSN network events. To perform this, we have derived a reinforcement learning model that learns $\mathbb{B}^*$ and ψ^* parameters online, without even estimating the unknown cost and TPF. To exhibit DPM and PHY switching decision learning the policy parameters is needed. Here, the derived network state learning model (NSLM) can be considered as a model-free reinforcement learning approach to obtain optimal state-action paid for DPM and power switching policy. In addition, Q-learning approach that measures $\mathbb{B}^*$ while assuming network parameters as unknown and varying is also developed in this paper. As an enhanced solution for traditional Q-Learning based network state learning (QNSL) to adopt mobility based WMSN, specifically for NCAM-RP [14] we have strengthened NSLM that could enable an integrated model comprising both the known as well as unknown network parameters. To achieve this, the buffer cost and the TCF are considered as learning variables. A brief of the conventional Q-Learning based DPM model is presented as follows:

3.4.4 Q--learningBased Network State Learning (QNSL)

The typical Q-learning scheme updates the step exhibited and evidences observed in individual time-slot on the basis of the evidence $\varsigma^{n} = (y^{n}, j^{n}, h^{n}, y^{n+1})$. mathematically,

$$\mathbb{B}^{n+1}(y^n, j^n) \leftarrow (1-\zeta^n)\mathbb{B}^n(y^n, j^n) + \zeta^n \left[h^n + \zeta \min_{a' \in \mathbb{A}} \mathbb{B}^n(y^{n+1}, j')\right. \tag{12}$$

In (12) y^{n} and ζ^{n} present the state and the performed action during each time slot n, respectively. The variableh^{n} presents the associated cost with related expectation value $h(y^{n}, j^{n})$ The results for next $(n+1)$ time slot is y^{n+1}, which is distributed according to the condition $m(y^{n+1} \mid y^{n}, j^{n})$. The parameter j' signifies the greedy action during states y^{n+1} which minimizes the present measurement-value of AVF. In (23), $\zeta^{n} \in [0,1]$ signifies a time-varying learning rate. Here, $\mathbb{B}^{0}(y,j)$ is initiated in random manner for the complete set $(y,j) \in \mathbb{Y} \times \mathbb{J}$.

To perform the approximation of $\mathbb{B}^*$, the traditional Q-learning model uses the average of the AVF $\mathbb{B}^{n}$. In practice, the fixed instantaneous cost and TPFs, iterative explore or visit the state-action pairs, and applying the stochastic approximation condition ζ^n, $\sum_{n=0}^{\infty}\zeta^n = \infty$ and $\sum_{n=0}^{\infty}(\zeta^n)^2 < \infty$, $\mathbb{B}^n$ converges with probability 1 to $\mathbb{B}^*$ as $n \to \infty$. Applying QNSL it can't be evident that what must be the optimal action to perform in each state during learning. Though, $\mathbb{B}^*$ can be obtained by random search of the present actions in each state; however, random search methods can't ensure optimal run-time performance because of iterative sub-optimal solution occurrences. On the other hand, the greedy actions exploring the information available in $\mathbb{B}^n$ can be significant to achieve certain expected level of performance. Undeniably, learning something already known can avoid the search process to get the optimal solution. Unlike other methods [15], in our model a method called ε-greedy action selection method is suggested. The Q-Learning assumption that the unknown cost and TPFs generally depend on the action is the main reason of the requirement of the search. Without hesitation, QNSL can't be stated as a robust and efficient approach due to the reason that it updates AVF for each pair of the state-action for complete time slots and does not achieve any known network parameter. Consequently, it suffers from convergence issue (high convergence time) to achieve the sub-optimal solution for DPM or PHY switching decision. This as a result can significantly reduce the run-time performance [17, 52]. Due to this reason, it can't be suggested for NCAM-RP based WMSNs, where time and bandwidth efficient transmission are must.

Considering such limitations, we have proposed NSLM is solved as MDP problem where the use of HMM and Lagrange relaxation makes it swift enough to achieve all major expectations.

3.5 Network State Learning Based DPM Model for NCAM-RP (NSLM) Implementation

Considering the practical NCAM-RP based WMSN scenario, where there are known as well as unknown network conditions, we developed NSLM for WMSNs dynamic power management and PHY switching scheduling. NSLM model targets achieving delay sensitive and bandwidth and energy efficient DPM along with optimal bandwidth utilization and minimum buffer cost, which are the inevitable needs for WMSNs. Unlike QNSL, which uses the sample average of the known network parameters for approximating the state transition parameters, NSLM uses both known as well as unknown network parameters to approximate network states for efficient DPM and PHY switching decision.

3.5.1 Partially Known Network State Parameters

Due to uncertain and varying topology NCAM-RP can have consistent topological variations and hence changing network status. Thus, there can be negligible or partial (i.e., very few) information available for run-time network functions. Before discussing the network state learning and scheduling, it is important to identify or define known and unknown parameters in WMSNs. Table 1 presents the parameter definition for WMSNs.

Table 1. Parameter Definition

	Deterministic	**Stochastic**
Known Parameters	PHY Power management (PPM) state, overall power cost	Good put, Holding cost
Unknown Parameters	-	Packet arrival distribution, CSI, overflow cost

Here, with known DPM state (interval (0,1)) the state transition can be stated to be stochastic in nature [50]. On the contrary, with unknown power management state the DPM transition can be stochastic as well as unknown. In same manner, if the BEP value is unknown for each time slot, then the respective good put and holding cost can be stated to be stochastic and unknown. The power transmission, the power cost can be considered as unknown. NSLM exploits the known as well as unknown parameters to enable swift PHY switching and eventual DPM for robust WMSN routing protocol.

3.5.2 NSLM Learning

In case of NSLM model, initially the network state is defined as soon as the known network dynamics have occurred just before the unknown network dynamics. Let the network state be $\tilde{y} \in \mathcal{Y}$, for state prediction and decision process NSLM at nth time slot is required to be associated with the current network state parameters given as $y^n = (q^n, c^n, p^n)$ and the action parameter $j^n = (\alpha^n, s^n, o^n)$. With these conditions, states at n^{th} and $(n+1)^{th}$ time slot would be (13) and (14) respectively.

$$\tilde{y}^n = (\tilde{q}^n, \tilde{c}^n, \tilde{p}^n) = ([q^n - f^n], c^n, p^{n+1}) \tag{13}$$

$$y^{n+1} = (q^{n+1}, c^{n+1}, p^{n+1}) = ([q^n - f^n] + \eta^n, c^{n+1}, p^{n+1}) \tag{14}$$

In above expression, the transmission buffer $\tilde{q}^n = q^n - f^n$ presents the state of the transmission buffer after packet transmission. NSLM applies known network parameters, especially the state transition information from y^n to y^{n+1} state after applying action p^n. In addition, the impending state applies all unknown network dynamics i.e. the number of packet arrivals f^n and next channel state information c^{n+1}. The buffer state at $(n+1)$ can be stated at time slot n $(\text{i.e.}, q^{n+1} = \tilde{q}^n + \eta^n)$. The relationship between different state parameters and associated actions are presented in Figure 3.

Observing Figure 3. it can be easily found that obtaining single DPM information may be significant to extract information about various other state-action pairs causing it. It puts foundation of the NSLM to exhibit efficient network state learning and transmission scheduling. The proposed learning model encompasses two types of the TPF components, known and unknown components.

The overall implementation model of the proposed NSLM model for DPM and PHY switching is briefed as follows:

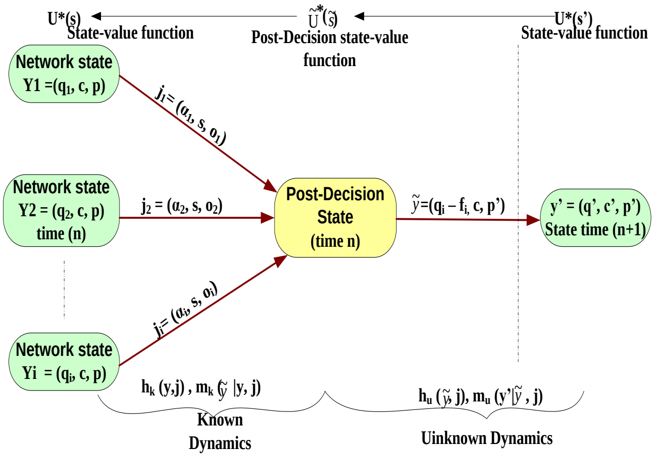

Figure 3. The State-Action relationship $(y_1, j_1), \ldots, (y_i, j_i)$

3.5.3 NSLM based optimal DPM policy

The proposed NSLM model converges to the optimal post decision state value function $\tilde{U}^{*,\mu}(\tilde{y})$ estimation even under dynamic or uncertain network condition, providing the learning rates α^n fulfills the following conditions:

$$\sum_{n=0}^{\infty} \zeta^n = \infty \text{ and } \sum_{n=0}^{\infty} (\zeta^n)^2 < \infty. \quad (15)$$

A brief of the post decision state learning based on NSLM for DPM and PHY switching decision is presented as follows:

Phase-1 Initialization: Initiate NSLM state value function $\tilde{U}^0$ at $n = 0$

Phase-2 Initiate Greedy Action: Initiate greedy action function

$$j^n = \arg\min_{j \in \mathbb{J}} \left\{ h_k(y^n, j) + \sum_{\tilde{y}} m_k(\tilde{y} \mid y^n, j) \tilde{U}^n(\tilde{y}) \right\} \quad (16)$$

Phase-3 Measure Network states, activities and experiences

$$\xi^n = (y^n, j^n, \tilde{y}^n, h_u^n, y^{n+1}) \quad (17)$$

Phase-4 Estimate the network state value y^{n+1}:

$$U^n(y^{n+1}) = \min_{j \in \mathbb{J}} \left\{ h_k(y^{n+1}, j) + \sum_{\tilde{y}} m_k(\tilde{y} \mid y^{n+1}, j) \tilde{U}^n(\tilde{y}) \right\} \quad (18)$$

Phase-5 Update NSLM value: With obtained information (Phase 3 and Phase-4) update the value of NSLM

$$\tilde{U}^{n+1}(\tilde{y}^n) \leftarrow (1 - \zeta^n)\tilde{U}^n(\tilde{y}^n) + \zeta^n[h_u^n + \zeta U^n(y^{n+1})] \quad (19)$$

Phase-6 Apply Lagrange multiplier to update decision parameters μ using equation (20)

Phase-7 Execute Iterations: Execute iteration for the functions and update time index $(n \leftarrow n + 1)$ and move to the Phase-2.

To enable delay sensitive, energy efficient and resource efficient transmission scheduling or PHY switching, NSLM applies a stochastic sub-gradient learning approach for estimating optimal Lagrange multiplier to be used in (11). It is obtained as (20).

$$\mu^{n+1} = \lambda[\mu^n + a^n(d^n - (1 - \rho)\vartheta)], \quad (20)$$

where λ assigns μ over $[0, \mu_{max}]$, a^n presents a time-dependent learning rate with similar characteristics as ζ^n. The variable d^n presents the transmission buffer cost having expectation $d(y^n, j^n)$, and $(1 - \rho)\vartheta$ converts the decremented delay constraint ϑ into an average delay constraint. In order to achieve swift convergence of (20) an additional condition is needed to be meet by a^n and ζ^n:

$$\sum_{n=0}^{\infty} (\zeta^n + a^n) < \infty \text{ and } \lim_{n \to \infty} \frac{a^n}{\zeta^n} \to 0. \quad (21)$$

Since, the conceptual convergence of NSLM requires a fixed Markovian condition, and therefore it becomes possible to track the Markovian dynamics by altering the learning rates ζ^n sequences (to update network state value function) and a^n (for Lagrange multiplier update). Maintaining ζ^n and a^n bounded away from zero can avoid previous evidences or the experience from biasing the state value function and the Lagrange multiplier. It can significantly enable the tracking of the network state dynamics. This mechanism in conjunction with the proposed NSLM learning can enable swift Markovian convergence to enable fast power management decision and transmission scheduling. This as a result assure optimal bandwidth utilization, minimal delay and energy efficient communication. These all achievements are the dominating needs for WMSN communication systems. Thus using the known network condition or states through NCAM-RP, the cost and transmission probably functions, $(h_k(y^n, j)$ and $m_k(\tilde{y}^n \mid y^n, j))$ our proposed learning model can perform efficient and timely network learning and transmission decision to achieve optimal DPM and PHY switching. Furthermore, the network state variations in mobility based WMSN could give rise to the unknown network dynamics $(h_u(\tilde{y}^n)$ and $m_u(y^{n+1} + \tilde{y}^n))$, which is independent of the execution function and therefore there is no need to use randomized search so as to identify the optimal action in the individual network state. It shows that the current or the latest estimates of the value function can be obtained and therefore there is no need to investigate or use any non-greedy action.

4. Results and Discussion

In last few years, the exponential rise in communication systems, particularly multimedia data communication over sensor networks have demanded efficient and robust routing approaches to ensure QoS delivery and energy efficiency. WMSN has emerged as one of the most demanding techniques to meet major low-cost communication purposes. To meet the key demands such as timely data delivery, higher throughput, data sensitive resource allocation and mission critical communication in our previous work [14], a cross layer architecture-based network condition aware routing protocol for WMSN named NCAM-RP was developed. Our effort exhibited optimal performance to meet WMSN requirements, however it could not deal with energy efficient and bandwidth efficient transmission. Since, these key factors are closely related to the PHY layer of the protocol stack enhancing dynamic power management and PHY switching was must. On the other hand, introducing mobility for WMSN requires an efficient approach to make optimal DPM decision under varying network conditions. With this motivation, in this research we hypothesized that learning network states dynamically and scheduling transmission accordingly may enable optimal resource utilization that not only enhance bandwidth utilization but can also enable time and energy efficient communication over WMSNs. With this motivation, in this paper, we have developed a novel dynamic network state learning based PHY scheduling and DPM for NCAM-RP. Since, our proposed NCAM-RP model employs mobile nodes to perform communication, it forces the system to undergo significantly higher topological variations and network state transition. This as a result makes transmission decision tough to ensure contention free, energy efficient and QoS enriched communication.

In this work, the DPM and network adaptive transmission decision has been considered as a problem for constrained MDP. The proposed NSLM functions as a stochastic approach that learns both the known as well as unknown network parameters to decide optimal transmission decision and PHY switching. The proposed NSLM model emphasizes on enabling an efficient dynamic power management and PHY switching at the physical layer of the protocol stack. The proposed model is developed in a way that it supports quadrature amplitude modulation (QAM) constellation in which gray cod can be used for mapping the information bits in QAM symbols. Being mobility based WMSN; it undergoes dynamic network or topological variations. In this case, there can be certain known as well as unknown network state or variables. Unlike known network variables, parameters such as packet receiving rate, bits error probability, channel status, buffer availability etc. could be the unknown parameters. Unlike traditional learning approach such as Q-Learning, which considers only unknown parameters as input, in our model we have incorporated both the known as well as unknown network state variables as input. In this paper, we have formulated dynamic network state learning model and DPM process in the form of a Markov decision process. The proposed model exploits the state value functions and associated to learn the network. Moreover, we derived buffer cost and transition probability function to learn the network states and derive optimal DPM and PHY switching policy. The proposed scheme was developed in a time slotted model, where for each time slot the network conditions or network states were obtained for further learning and prediction. We have exploited the Lagrange multiplier factor to enable swift convergence of the proposed learning model to achieve more time efficient DPM and transmission scheduling. The overall proposed simulation model is developed using MATLAB 2015a software tool.

To perform simulation, we have applied the decremental factor (DF) as $\gamma = 0.98$ that enables NSDL to converge fast. This as a result has strengthened NSDL to perform fast PHY switching decision and DPM. This is must for multimedia communication over sensor networks to ensure QoS delivery. To estimate the signal to noise ratio (SNR) or RSSI the power, noise and bandwidth have been applied. Here, SNR is obtained by (22).

$$SNR = \frac{P^0{}_{Tx}}{N_0 W} \qquad (22)$$

where $P^0{}_{Tx}$ presents the transmission power each slot n, the noise component in the proposed learning model is given by N_0, and the bandwidth is presented by term W. In our simulation, the bandwidth is considered as

$$1/T_s \qquad (23)$$

where T_s presents the symbol period.

Some of the key simulation parameters applied in our research model is given in Table 2.

Table 2. Simulation parameters

Parameter	Simulation Value
Packet arrival rate (Avg.)	220 packet/sec
Bits per symbol (BPS)	[1,2,3,…,10]
Buffer Size	25 data packets
Channel State	[-18.8, -13.8, -11.2, -9.3, -7.8, -6.3, -4.7, -2.1]
Decremental of discount factor	0.98
Holding cost	4 packets
Noise PSD	2×10^{-11} watts/Hz.
Energy/Power in OFF status	0 Watts
Energy/Power in ON status	80/160.320 watts
Size of Packet	5000 bits
DPM or PHY Switching Actions	P_{ON}, P_{OFF},
Power Management status	ON, OFF
Symbol Rate	5.0×10^5 symbols/sec.
Time slot	10 ms.
Dynamic Traffic (Tx) variation	Dynamic traffic load (0,1,2,3,…,10) packets/sec.

Unlike Q-Learning based DPM, where sample average of the AVFs are used to approximate Q^* parameter, our proposed NSLM model exhibits network learning because of the sample average value of the post decision state functions and associated variables like, resource states, bit error probability, power states etc. It enables the approximation of $\tilde{V}^*$ swift and thus making NSLM time, energy as well as bandwidth efficient. To assess the performance of our proposed NSLM based DPM and PHY switching model for WMSN, the overall proposed model is examined in terms of buffer cost, bandwidth utilization, holding cost, buffer overflow conditions, etc. To further examine the performance of NSLM approach, a comparative model using Q-Learning is developed for DPM. NSLM applies both the known as well as unknown network parameters, while Q-Learning applied only unknown parameters to perform DPM. Here, we considered Q-Learning because of it model free reinforcement learning ability to obtain action-selection measure for MDP.

The buffer cost, which can be stated analogous to the time during which the data traverses across buffer does impact the overall performance of WMSN, especially for timely data delivery. To have QoS communication over multimedia sensor network or WMSN, ensuring timely delivery is must and therefore minimal buffer cost is anticipated to avoid jittering type problems. Considering the results obtained, here it can be found (Figure 4.) that the proposed NSLM exhibits significantly lower buffer cost than its counterpart Q-learning based DPM. Here, it should be noted that the prime use of buffer cost was to reduce the queuing delay in the transmission buffer to enable swift data transmission. This as a result avoids the buffer overflow problems that could lead significantly higher transmission delay, and thus degrading QoS delivery of WMSNs. In the proposed DPM model, buffer cost (we stated it as cost in Figure 4.) has been considered as the sum of holding cost and overflow cost with respect to the rate of data arrival and good put distribution. Here, the holding cost signifies the duration for which the data remains in the buffer. Observing Figure 5. it can be easily found that the proposed NSLM model reduces holding cost significantly and thus enabling lower buffer cost as depicted in Figure 4. The prime reason behind such optimization could be the consideration of both the known as well as unknown network states to enable decision state function for DPM and adaptive power switching.

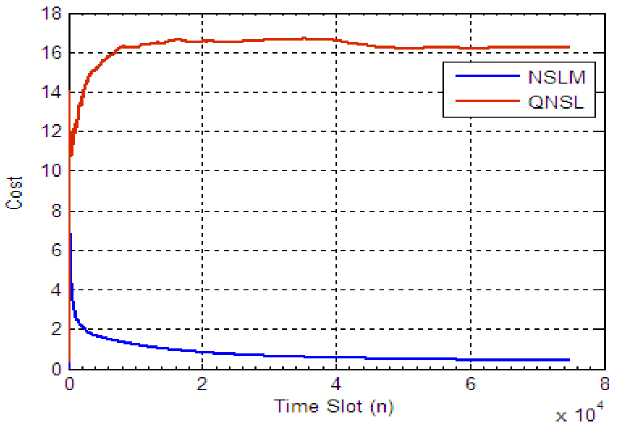

Figure 4. Buffer cost analysis

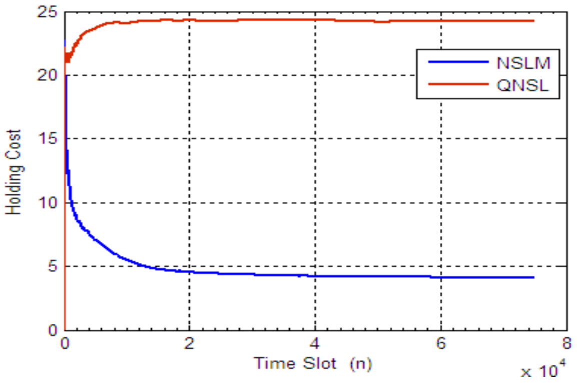

Figure 5. Holding cost analysis

As already stated that the presented work intends to contribute a novel dynamic network state learning approach for NCAM-RP functional over WMSN systems, it is inevitable to consider burst type transmission or packet arrival in real time environment for performance analysis. The burst type data arrival can have the direct impact on network stability, data drop, data retransmission, and delay. Considering it as vital factor for performance assessment, we have examined NSLM model as well as Q-learning based network state learning (QNSL) model and the comparative result obtained is presented in Figure 6. From the obtained results (Figure 6.), it is confirmed that the proposed approach exhibits significantly better overflow avoidance than QNSL. In directly, the efficacy of overflow avoidance strategy is directly related to adaptive data transmission and buffer utilization. Better buffer utilization can assure minimum overflow. This hypothesis is also confirmed by obtained results (Figure 7 and Figure 8)

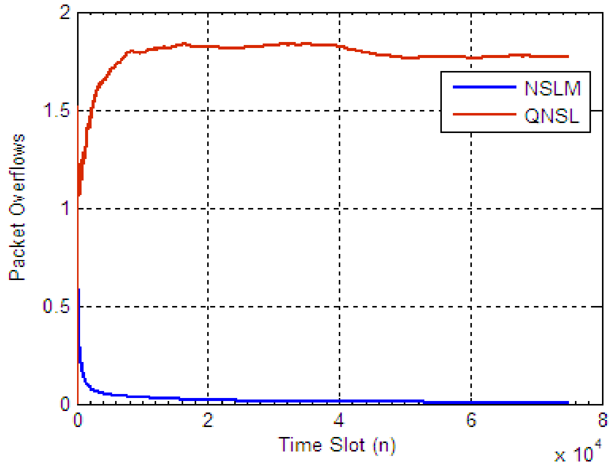

Figure 6. Packet overflow analysis

One of the key QoS centric requirements of WMSNs is efficient bandwidth utilization that eventually leads optimal data transmission and its timely delivery to assure quality of experience (QoE) or perception. Being a post decision type reinforcement learning model NSLM employs dynamic buffer conditions too to achieve optimal DPM policy and transmission decision. In fact, the bandwidth utilization by any transmission model or PHY switching function can be visualized through its transmission activity. Better transmission activity can lead better resource utilization. This hypothesis can be easily observed in Figure 7 and Figure 8. Observing the results in Figure 7 and Figure 8, it can be easily found that NSLM exhibits better transmission that as a result enables higher resource utilization. On the other hand, as depicted in Figure 7, QNSL performs relatively less efficient or poor packet transmission that results into lower bandwidth occupancy. It exhibited limited efficiency of the QNSL approach. Considering other parameter named cumulative average cost (CAC) which is close related to the overall DPM process and transmission mechanism, it can be found that the CAC is minimum for NSLM based DPM as compared to the QNSL model (Figure 7 and Figure 8).

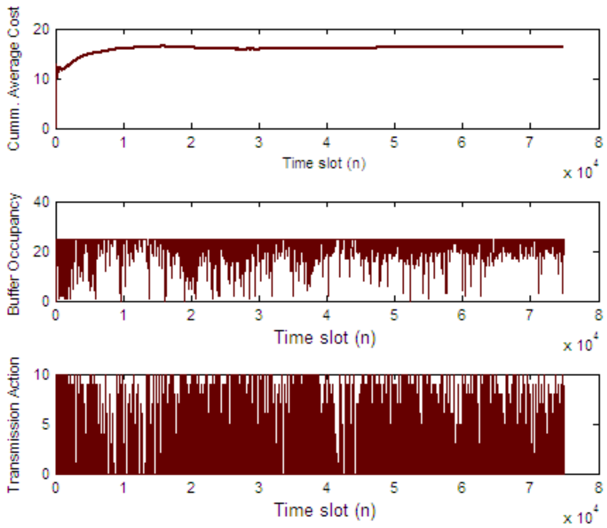

Figure 7. Buffer occupancy and transmission action using Q-Learning based DPM

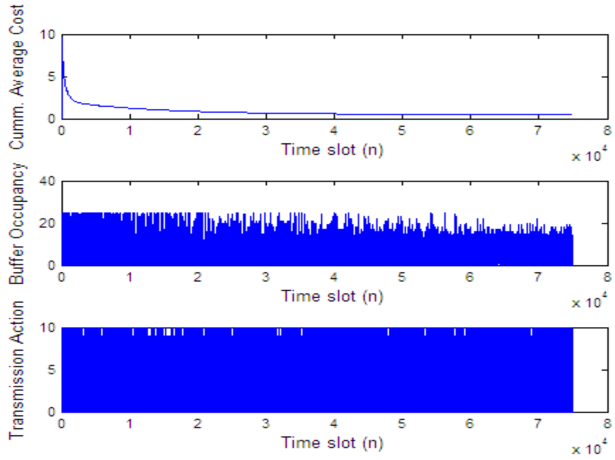

Figure 8. Buffer occupancy and transmission action using Proposed NSLM based DPM

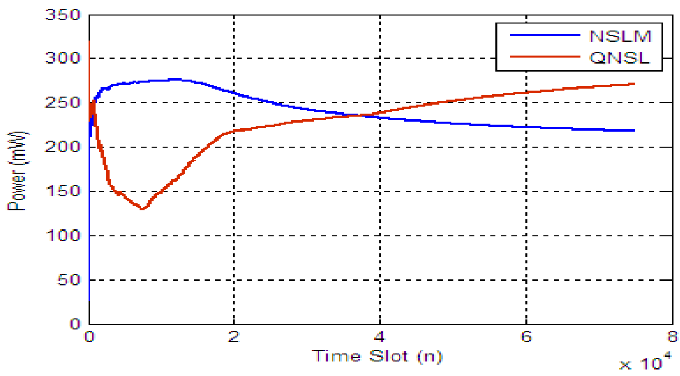

Figure 9. Power consumption during transmission

Thus, the overall results for bandwidth efficiency state that the proposed NSLM based DPM model can be of paramount significance for multimedia transmission over WMSN under uncertain or dynamic topology scenario. This is the matter of fact that energy efficiency is one of the decisive performance parameters of any sensor network or communication system. Enabling lower energy or power consumption can make transmission system more efficient, particularly to enhance network life time. Figure 9 exhibits that the proposed NSLM model performs timely data delivery with optimal bandwidth utilization, while ensuring minimal energy consumption. Here, NSLM can be found more energy efficient than the QNSL based DPM and transmission approach for WMSN.

5. Conclusions

The demands of multimedia communication have established WMSNs to be one of the most potential communication network to serve real time applications, including multimedia broadcast, security and surveillance systems, intelligent transport system, industrial automation etc. These all applications demand certain efficient routing approach to ensure timely, energy and bandwidth efficient communication. On the other hand, mobility based WMSNs can enable an array of applications serving timely and low-cost communications. However, mobility eventually results into significantly higher topological variations and network uncertainty. In such cases managing optimal physical layer switching or transmission switching while avoiding contention and packet loss becomes highly intricate. Furthermore, for WMSNs, assuring higher throughput, minimum delay, optimal bandwidth utilization, and energy efficiency are the key requirements. Enabling an efficient dynamic power management and PHY switching policy has been considered as an optimistic measure to achieve these objectives. Considering network state uncertainty during run-time (particularly with mobility), in this paper a novel reinforcement learning approach called network state learning model (NSLM) based stochastic optimization model has been developed that considers known as well as unknown network states to derive an optimal state action pairs for dynamic power management policy. Unlike traditional learning approach such as Q-Learning, we have incorporated both the known as well as unknown network parameters to perform transmission decision, where the DPM is considered as a problem for constrained Markov decision process (MDP). In the proposed model, two functions, cost functions and the network state transition probability functions are used as input variables to learn the network and derive optimal transmission policy. Here, the use of Hidden Markov Model (HMM) and Lagrange multiplier-based approach has enabled swift convergence that as a result makes transmission decision fast. The overall results obtained exhibit that the proposed NSLM based DPM and transmission scheduling achieves optimal bandwidth utilization, low power consumption, significantly low buffer cost, holding cost and overflow. These enhanced outcomes confirm suitability of the proposed DPM and transmission scheduling policy to enable efficient WMSNs solution. In future, the efficacy of other stochastic optimization approaches can be explored.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “A survey onwireless multimedia sensor networks,” Comput. Net, Vol. 51, no. 4, pp. 921_60, Mar. 2007.

[2] D. Guinard, V. Trifa, T. Pham, and O. Liechti, “Towards Physical Mashups in the Web of Things,” in Proc. INSS’09, pp. 196_99, Jun. 2009.

[3] P. F. James, C. How, D. Bull, and A. Nix, “Distortion-based link adaptation for wireless video transmission,” EURASIP J. Adv. Signal Process., Vol. 2008, pp. 1_17, Mar. 2008.

[4] M. V. Schaar, and D. S. Turaga, “Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission,” IEEE Trans. Multimedia., Vol. 9, no. 1, pp. 185_197, Jan. 2007.

[5] Y. P. Fallah, P. Nasiopoulos, and H. Alnuweiri, “Efficient transmission of H.264 video over multirate IEEE 802.11e WLANs,” EURASIP J. Wireless Commun. Net, Vol. 2008, pp. 1_14, Mar. 2008.

[6] T. Melodia, and I. F. Akyildiz, “Cross-layer QoS-Aware communication for ultra wide band wireless multimedia sensor networks,” IEEE J. Select. Areas Commun., Vol. 28, no. 5, pp. 653_63, Jun. 2010.

[7] L. Shu, M. Hauswirth, L.Wang, Y. Zhang, and J. H. Park, “Cross-layer optimized data gathering in wireless multimedia sensor networks,” in Proc. CSE’09, pp. 961_66, Oct. 2009.

[8] [8] F. Oldewurtel, J. Ansari, and P. Mahonen, “Cross-layer design for distributed source coding in wireless sensor networks,” in Proc. SENSORCOM’ 08, pp. 435_43, Aug. 2008.

[9] L. Zhou, X. Wang, W. Tu, G. Mutean, and B. Geller, “Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks,”

IEEE J. Select. Areas Commun., vol. 28, no. 3, pp. 409_19, April 2010.
[10] S. Li, R. Neelisetti, C. Liu, S. Kulkarni, and A. Lim, “An interference aware routing algorithm for multimedia streaming over wireless sensor networks,” Int. J. Multimedia Appl., vol. 2, no. 1, pp. 10_30, Feb. 2010.
[11] N. Saxena, A. Roy, and J. Shin, “Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks,” Comput. Net, vol. 52, no. 13, pp. 2532_42, Sept. 2008.
[12] A. Jin, W. Song, P. Ju, D. Zhou, "Energy-Aware Cooperation Strategy With Uncoordinated Group Relays for Delay-Sensitive Services,” in Vehicular Technology, IEEE Transactions, vol.63, no.5, pp.2104_14, Jun 2014.
[13] J. Sen, A. Ukil, "An adaptable and QoS-aware routing protocol for Wireless Sensor Networks," in Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 1st International Conference, pp.767_71, May 2009.
[14] A. Ajina, Mydhili, “Cross Layered Network Condition Aware Mobile Wireless Multimedia Sensor Network Routing Protocol for Mission Critical Communication,” in International Journal of Communication Networks and Information Security (IJCNIS), pp.48-56, Vol. 9, no. 1,pp. 48-56, April 2017.
[15] R. S. Sutton and A. G. Barto, “Reinforcement learning: an introduction,” Cambridge, MA:MIT press, 1998.
[16] N. Mastronarde, and M. van der Schaar, “Fast Reinforcement Learning for Energy-Efficient Wireless Communication,” in IEEE Transactions on Signal Processing, vol. 59, no. 12, pp. 6262_66, Dec. 2011.
[17] E. Evan-Dar and Y. Mansour, “Learning rates for Q-learning,” Journal of Machine Learning Research, vol. 5, pp. 1_25, Dec. 2003.
[18] M. Triki, A. C. Ammari, Y. Wang, and M. Pedram, “Reinforcement learning algorithms for dynamic power management,” 2014 World Symposium on Computer Applications & Research (WSCAR), Sousse, pp. 1_6, Jan. 2014.
[19] H. Shen, Y. Tan, J. Lu, Q. Wu, and Q. Qiu, “Achieving Autonomous Power Management Using Reinforcement Learning,” ACM Transactions on Design Automation of Electronic Systems, Vol. 18, no. 2, pp. 1_32, Mar. 2013.
[20] V. A. Shiny, V. Nagarajan, “Energy Efficient Routing Protocol for Mobile Wireless Sensor Network,” Int. Journal of Computer Applications, Vol. 43, no.21, pp. 1_5, April 2012.
[21] K. Khanke, M. Sarde, “An Energy Efficient and QoS Aware Routing Protocol for Wireless Sensor Network,” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, no. 7, pp. 1_4, July 2015.
[22] S. Zhuo, Z. Wang, Y. Song, Z. Wang, and L. Almeida, “iQueue-MAC: A Traffic Adaptive duty-cycled MAC Protocol with Dynamic Slot Allocation,” IEEE International Conference on Sensing, Communications and Networking (SECON), pp. 95_103, June 2013.
[23] L. Zhang, and Y. Zhang, "Energy-Efficient Cross-Layer Protocol of Channel-Aware Geographic-Informed Forwarding in Wireless Sensor Networks," in IEEE Transactions on Vehicular Technology, vol. 58, no. 6, pp. 3041-3052, July 2009.
[24] A. Alhalafi, L. Sboui, R. Naous, and B. Shihada, "gTBS: A green Task-Based Sensing for energy efficient Wireless Sensor Networks," 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), San Francisco, CA, pp. 136-143, April 2016.
[25] M. Al-Jemeli, and F. A. Hussin, "An Energy Efficient Cross-Layer Network Operation Model for IEEE 802.15.4-Based Mobile Wireless Sensor Networks," in IEEE Sensors Journal, vol. 15, no. 2, pp. 684_92, Feb. 2015.
[26] Y. Chen, L. Gao, Y. Xing, and W. Yi, "Cross-Layer Design for Energy-Efficient Reliable Routing in Wireless Sensor Networks," 2015 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN), Shenzhen, pp. 31_6, Dec. 2015.
[27] A. Shareef, A. Shareef, and Y. Zhu, "Optrix: Energy aware cross layer routing using convex optimization in wireless sensor networks," 2015 IEEE International Conference on Networking, Architecture and Storage (NAS), Boston, MA, pp. 141_50, Aug. 2015.
[28] T. Jaradat, D. Benhaddou, M. Balakrishnan, and A. Al-Fuqaha, "Energy efficient cross-layer routing protocol in Wireless Sensor Networks based on fuzzy logic," 2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, pp. 177_82, July 2013.
[29] G. S. Brar, S. Rani, V. Chopra, R. Malhotra, H. Song, and S. H. Ahmed, "Energy Efficient Direction-Based PDORP Routing Protocol for WSN," in IEEE Access, vol. 4, pp. 3182_94, Jan. 2016.
[30] A. Tanwer, M. Hussain, and P. S. Reel, "Energy efficient wireless sensing framework to enhance mobile learning," 2010 International Conf.on Tech. for Education, Mumbai, pp. 50_5, July 2010.
[31] Walid Osamy and Ahmed M. Khedr, “An algorithm for enhancing coverage and network lifetime in cluster-based Wireless Sensor Networks”, International Journal of Communication Networks and Information Security (IJCNIS), Vol. 10, No. 1, pp. 1-9, April 2018.
[32] M. Chen, V. C. Leung, S. Mao, and Y. Yuan, “Directional geographical routing for real-time video communications in wireless sensor networks,” Comput. Commun., vol. 30, no. 17, pp. 3368_83, Nov. 2007.
[33] H. P. Shiang, and M. Var der Schaar, “Distributed resource management in multi-hop cognitive radio networks for delay sensitive transmission,” IEEE Trans. Veh. Technol., vol. 58, no. 2, pp. 941_53, Oct. 2009.
[34] P. Frossard, J. C. De Martin, and R. Civanlar, “Media streaming with network diversity,” Proc. IEEE, vol. 96, no. 1, pp. 39–53, Jan. 2008.
[35] J. B. Othman, and B. Yahya, “Energy efficient and QoS based routing protocol for wireless sensor networks,” J. Parallel Distrib. Comput., vol. 70, no. 8, pp. 849_57, Aug. 2010.
[36] E. Felemban, S. Member, C. Lee, and E. Ekici, “MMSPEED:Multipath multi-SPEED protocol for QoS guarantee of reliability and timeliness in wireless sensor networks,” IEEE Trans. Mobile Comput., vol. 5, no. 6, pp. 738_54, June 2006.
[37] “Amendment : Medium Access Control (MAC) Quality of Service (QoS) Enhancements,” IEEE Std. 802.11e, Jul. 2005.
[38] H. Kimand, and S. G. Min, “Priority-based QoSMAC protocol for wireless sensor networks,” in Proc. IEEE IPDPS’09, pp. 1_8, Nov. 2009.
[39] S. Ehsan, B. Hamdaoui, "A Survey on Energy-Efficient Routing Techniques with QoS Assurances for Wireless Multimedia Sensor Networks," in Communications Surveys & Tutorials, IEEE, vol.14, no.2, pp.265_78, Jan. 2012.
[40] B. Yahya, J. Ben-Othman, "An energy efficient and QoS aware multipath routing protocol for wireless sensor networks," in Local Computer Networks, IEEE 34th Conference, pp.93_100, Oct. 2009.
[41] Wang Xiao-ming, Li Cheng-bo, Lu Jun-ling, Liu Ding, "QoSPMR: QoS-Aware and Priority-Based Multipath Routing Algorithm for WMSNs," in Engineering and Technology (S-CET), pp.1_5, May 2012.
[42] Syarifah Ezdiani Syed Nor Azlan, Adnan Al-Anbuky, “Modelling the Integrated QoS for Wireless Sensor Networks

with Heterogeneous Data Traffic," Open Journal of Internet Of Things (OJIOT), Vol. 1, no. 1, pp. 1_15, 2015.
[43] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," IEEE Trans. on Information Theory, vol. 50, no. 1, pp. 1_21, Jan. 2004.
[44] R. Berry, and R. G. Gallager, "Communications over fading channels with delay constraints," IEEE Trans. Inf. Theory, vol 48, no. 5, pp. 1135_49, May 2002.
[45] N. Salodkar, A. Bhorkar, A. Karandikar, and V. S. Borkar, "An on-line learning algorithm for energy efficient delay constrained scheduling over a fading channel," IEEE Journal on Selected Areas in Communications, vol. 26, no. 4, pp. 732_42, Apr. 2008.
[46] M. H. Ngo, and V. Krishnamurthy, "Monotonocity of constrained optimal transmission policies in correlated fading channels with ARQ," IEEE Trans. on Signal Processing, vol. 58, no. 1, pp. 438_51, Jan. 2010.
[47] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design," IEEE Trans. on Wireless Communications, vol. 4, no. 3, pp. 1142_53, May 2005.
[48] D. Krishnaswamy, "Network-assisted link adaptation with power control and channel reassignment in wireless networks," in Proc. 3G Wireless Conf., pp. 165_70, Dec. 2002.
[49] K. B. Song, and S. A. Mujtaba, "On the code-diversity performance of bit-interleaved coded OFDM in frequency-selective fading channels," in Proc. IEEE Veh. Technol. Conf., vol. 1, pp. 572_76, Oct. 2003.
[50] L. Benini, A. Bogliolo, G. A. Paleologo, and G. De Micheli, "Policy optimization for dynamic power management," IEEE Trans. on computer-aided design of integrated circuits, vol. 18, no. 6, pp. 813_33, June 1999.
[51] D. Bertsekas, and R. Gallager, "Data networks," Prentice Hall, Inc., Upper Saddle River, NJ, 1987.
[52] P. Auer, T. Jaksch, and R. Ortner, "Near-optimal regret bounds for reinforcement learning," Advances in Neural Information Processing Systems, vol. 21, pp. 89_96, Nov. 2009.