An Optimal Management Modelling of Energy Harvesting and Transfer for IoT-based RF-enabled Sensor Networks

Article *in* DSVL Detecting Selfish Node In Vehicular Ad-hoc Networks (VANET) by Learning Automata · January 2020

Pradeep Kumar T S
VIT University Chennai

Rafidah Md. Noor
University of Malaya

Ismail Ahmedy
University of Malaya

Saaidal Razalli
University of Malaya

An Optimal Management Modelling of Energy Harvesting and Transfer for IoT-based RF-enabled Sensor Networks

SHAIK SHABANA ANJUM[1*], RAFIDAH MD NOOR[2†], ISMAIL AHMEDY[3], MOHAMMAD HOSSEIN ANISI[4], SAAIDAL RAZALLI AZZUHRI[5], MISS LAIHA MAT KIAH[6], JAIME LLORET[7], PRADEEP KUMAR[8]

[1,2,3,5,6]*Department of Computer System and Technology, Faculty of Computer Science and Information technology, University of Malaya, 50603,Malaysia.*
[1,2]*Centre for Mobile Cloud Computing Research (C4MCCR),Faculty of Computer Science and Information technology, University of Malaya, 50603,Malaysia.*
[4]*School of Computer Science and Electronic Engineering, University of Essex,Colchester,United Kingdom.*
[7]*Integrated Management Coastal Research Institute,Universitat Politcnica de Valncia,Spain.*
[8]*School of Computing Science and Engineering,Vellore Institute of Technology,Chennai,India.*

A crucial conduct norm for a sensor network is to avoid network failures and packet drop. One of the other essential requirements is to effectively manage the energy levels of the nodes according to the states of the operation required for an application. This paper focuses to propose an energy management model with the aim of allowing energy optimization of Radio Frequency(RF)-enabled Sensor Networks (RSN) during the process of Energy Harvesting (EH) and Energy Transfer (ET) through controlled optimization. Primarily, energy harvesting of sensor networks through RF signals is focussed in this research to address the drawback of frequent replacement of batteries, persistent recharge request, dead state of nodes and periodical eradication of bat-

* email: shabana@um.edu.my
† email: fidah@um.edu.my

teries. Secondly, this paper focuses on mathematical modelling of the RF sensor nodes within the proposed Energy Harvesting RSN (EHRSN) and Energy Transfer RSN (ETRSN) framework of Energy Management RSN model (EMRSN) where the nodes are characterized as Semi Markov Decision Process (SMDP) and optimal policies are computed for numerically evaluating and analysing the issue of higher energy consumption. The most optimal state transitions are computed and mathematically formulated based upon stochastic dynamic programming to carry out the numerical analysis. It has been found that through controlled optimization, the sensor networks when energized through RF energy for EH process, the probability of 0.8 or more works best at the lower power level. On the other hand, for ET, the sensors tend to work more when the probability is either 0.8 or more at higher power levels. The results obtained are further employed to program the sensors accordingly in the Internet of Things (IoT) contexts during EH and ET processes to achieve maximum throughput, network lifetime and energy efficiency.

Key words: Internet of Things, Energy Management, Energy harvesting, Energy Transmission, Energy Optimization.

1 INTRODUCTION AND MOTIVATION

Wireless Sensor Networks (WSN) comprises of varied applications such as medical, health-care, military surveillance, habitat monitoring and industrial control. The functionality of these applications is to transmit the sensed information over dedicated communication channels and networks provided by the IoT technology. This technology enables the respective sensors to record and reciprocate with the sensing information through a layered architecture. The data delivery process carried out in an efficient and reliable manner plays the key role in the performance of sensor networks for smarter IoT contexts [1]. An integrated Internet of Things (IoT) system comprises of sensor nodes that suffer drawbacks such as limited lifetime, area coverage, smaller size and limited processing power. Each sensor node is designed with sensing and data processing components for communicating with its neighbour nodes over dedicated channels. The node transfer data either directly or through relays to the sink node, which is forwarded to the end user or server via a gateway and the internet. Hence, these wireless class of networks poses the challenge of centralized data forwarding, management, control of sink node's placement and area coverage of access points in an IoT domain. In contrary to the IoT networking devices, which have IP address-based approach, the sensor nodes work based upon node ID or index addressing. Moreover, the energy consumption in sensor networks is inversely proportional to the mesh formation of the nodes. The feature of mesh formation by the sensor nodes can hence play a vital role to improve the lifetime and sustainability of sensor networks.

The energy optimization and management are handled by the MAC layer in sensor networks, whereas for IoT devices, the energy optimization is hardly possible since IP based access consume a huge amount of power across all the layers of the stack. Such humongous energy consumption is handled by IPv6 over Low-power Wireless Personal Area Networks (6LoWPAN) operating in the 2.4 GHz range of frequency with 250Kbps of data transfer rate. The IPv6 that is responsible for the networking and communication of switched packets and datagram transmission aids in the mitigation of power consumption during compression of these packets over the data link layer. The low-power WSN communications based on IEEE 802.15.4 and 6LoWPAN contribute and support actuating capabilities using low-power WSN devices and capillary communications for applications that need the sensing operations. Therefore, in a nutshell, it can be said that WSN forms to be a subset of a larger IoT domain. Since devices in both the technologies rely on energy storage devices for their functioning, it is necessary to focus on the aspect of energy optimization. IoT research has the capability to encapsulate the identification potential, sensing technology, artificial intelligence, and interconnection of nano - things, ultimately striving towards the objective of developing seamlessly interoperable and securely integrated systems. These integrated communication networks comprise of many interconnected units such as processor, memory, energy storage unit, radio, micro-controller and so on. The energy consumed by these units is very high during communication. Therefore, optimization of this energy consumption is a primary necessity to increase the lifetime of integrated systems. This research focuses to enable energy optimization for different states of a sensor node when it is energized through RF signals and therefore lead to improved network performance and increased lifetime of nodes. The work contributed by [2] describes the EM issues and taxonomy out of which the authors intend to solve the problem of two main aspects of higher energy consumption of sensor nodes- Energy Harvesting (EH) and Energy Transfer (ET).

This article basically provides description about how sensor nodes can be modelled according to controlled optimization process so that energy con-

sumption is minimized during the active operation of a sensor. The formulations and optimal policy computation fetch the results of how a sensor node need to be modelled for lighter and heavier sensing operations in IoT applications.
The related work of this research has been described in section 2 covering the aspects of need for energy efficient protocols, description about clustering techniques and the energy modelling concepts. Furthermore, the energy management model for RF enabled sensor networks followed by insights into the concept of Markov decision process and Dynamic programming has been elaborated and explained in section 3. The process and methodology of controlled optimization mechanism for the proposed system has been elaborated in section 4 followed by explanation and illustrations on how the numerical evaluation is carried out for nodes during EH and ET processes. The numerical analysis and results have been described in section 5. The experimental findings along with graphical representations has been elaborated in section 6. Finally, the conclusion is included as section 7 which summarizes the prior sections and the significant outcome of the research along with the future work.

2 RELATED WORK

This section elaborates the related work and background about the research contributions for energy conservation and optimization in sensor networks. It also provides insights into how the current work differs with existing solutions. The need for energy efficient routing protocols is a necessity for increasing the lifetime of the network and to prevent node failure [3], [4]. The operation and maintenance of sensor nodes in an intimidating environment and also the presence of error-prone communication links expose these networks to low energy levels thereby hindering the overall network performance and throughput. The networked infrastructure of heterogeneous devices in IoT contexts are equipped with sensors, controlling processors, wireless transceivers, and energy resources for data transmission and monitoring activities [5]. One of the dominant hurdles for implementing such interoperable networks is supplying adequate energy for network operation without compromising on Quality of Service (QoS). Hence, it is important to improve the energy efficiency of the connected devices in the sensor network by giving importance to the factor of battery consumption and energy drain. Despite the fact that there are several energy efficient protocols that have been designed to prolong the lifetime of the sensor nodes in traditional WSN, the integration of mobility-enabled technology with conventional static sensor networks, promises a new solution that balances energy consumption among the sensor nodes and eventually extends the lifetime of the network. The previous research contributions from [6], [7] and [8] can be referred for further insights and clarity. The power management protocols can be implemented either as independent sleep/wake-up protocols running on top of a MAC protocol (typically at the network or application layer), or strictly integrated with the MAC protocol itself [9].

The major difference between the current research contribution and existing solutions is that the nodes are modelled according to the controlled optimization of semi-markov process. This stochastic modelling enhances the optimality and lifetime of a sensor node thereby reducing the energy wastage due to idle listening. A standard protocol needs to be used for communication in sensor networks. Emerging communication standards such as IEEE 802.15.4 is being used in wireless sensor networks as an underlying protocol for building other standardized communication protocols such as ZigBee and LowPAN [10]. A critical performance criterion in backscatter modulation-based RFID sensor networks is the distance at which an RFID reader can reliably communicate with passive RFID sensors (or tags). Researchers in [11] have proposed a mechanism to introduce a power amplifier (PA) and an energy storage device (such as a capacitor or a battery), in the hardware architecture of conventional passive RFID tags, with the aim of allowing amplification of the backscatter signal to increase the read range of Radio Frequency(RF)-enabled Sensor Networks (RSN) during communication. On the contrary to the traditional wireless class of networks, the criterion for RF-based EH network's routing protocol includes circuitry design of the nodes and propagation of RF energy factor. The reason for this difference is the distinctive amount of RF energy that can be harvested by the active nodes during each operational cycles of the network.

The EH parameters predominantly define the routing metric. These parameters include quality of network link, the sensitivity of RF energy harvester, the distance between the nodes and RF sources, number of available communication channels, hop count and on the rate of conversion for harvested RF energy. There are research contributions, which deal with the relationship and correlation between RF energy recharging and sensor network routing. The heterogeneity and movement of the nodes can be used to determine the congestion in a network using Priority-based application-specific congestion control clustering (PASCC) protocol, which integrates the mobility and heterogeneity of the nodes to detect congestion in a network [12].

The sensor nodes in a network comprised of two components - energy dischargers (transceivers, radio, antenna, connectors, sensors, tags, readers) and energy suppliers (capacitors, batteries and so on). The nodes are energized by the use of batteries or capacitors. The real-time batteries tend to discharge energy even when they are idle as when compared to the simulators of linear characteristics. A considerable amount of energy is lost for every charging and recharging cycles, thereby leading to lesser voltage retention and a complete failure mode of the battery. The solutions for efficient optimization with the utilization of empirical, abstract, and physical models have been suggested by [13]. In physical prototyping, electrochemical batteries were employed whose reactions lead to either charging or discharging of energy storage devices. The behaviour of these devices can well be predicted using stochastic and abstract models whereas empirical modelling employs mathematical equations for charging and recharging of batteries. Computational RFID (CRFID) runtime, namely dewdrop employs an exponentially adaptive polling interval for the purpose of gathering energy over longer ranges of input power and huge target voltages [14]. The system model as described by [15] studies the pattern of a single sensor node using Discrete-Time Markov Chain (DTMC) modeling technique. In DTMC, the transmission time for a data unit decides the time slotting technique, which means that the duration needed to transmit one unit of data along with the MAC layer overhead. A heterogenous two-tier WSNs comprising of two distinctive and hierarchical set of nodes: sensor-tier nodes (M) and processing-tier nodes (N) is studied by [16] to explore and resolve the coverage processes of the deployed region using optimization theory, power control and clustering techniques. The efficiency of a harvester circuit is iterative and is affected through energy losses from switching process of MOSFET and diode to inductive resistance and power consumption of the comparator [17].

The remaining energy in the nodes and the distance of transmission has been suggested by [18] for energy optimization of WSN with user-defined software notifications. These sensor nodes provide NP-hardness since it has optimal energy consumption. Moreover, when such control nodes are designed to perform multiple monitoring operations at the same time, it consumes more energy. This issue is overcome using Particle Swarm Optimization (PSO) technique. The modelling method is utilized for selection of either free space or multiple modes. This is the method for energy modelling using the distance of transmission. The type of propagation model is selected based upon the transmission distance and the threshold. The PSO technique models the distance between the control nodes and its neighbour to be shorter, so that the overall energy consumption of the network is smaller. In energy-efficient multi-sink clustering algorithm (EMCA) proposed by [19] the CH transmits the forwarded data through shortest distance towards the gateway sink. The energy expenditure that is caused because of longer transmissions can be mitigated through multiple hopping strategy. Another aspect to be considered for energy modelling is the network topology across the network. Clustering method has been utilized, suggested and explored by many researchers in the wireless networks research domain. It is utilized for addressing the energy limitations across larger density of application-specific nodes. The working of remotely deployed sensor networks is designed to be self- configured, therefore, clustering techniques adapt to the automatic grouping of sensors to monitor the rate of energy consumption. The concept behind clustering technique is that a Cluster Head (CH) is elected by the nodes to transmit the data packets. The CH acts as a relay node and performs data aggregation and forwarding to the gateway sink. The characteristic parameters for a CH are usually modelled based upon its distance to the sink node, the range of transmission and residual energy levels. Since all the reception and transmission operations are handled by the CH there is more energy consumption and communication overhead which may ultimately lead to packet drop, CH elimination and network performance degradation. The other energy modelling methods can be adopted using Markov Decision Process (MDP) where the problems are formulated and solved using Dynamic Programming (DP) and Reinforcement Learning (RL). These tools are utilized for solving stochastic and control optimization issues.

A detailed explanation of the energy modelling concepts adopted from MDP and DP for proposed scenario will be discussed in the subsequent sections of this paper. Basically, the performance evaluation of any algorithm/techniques is done via three methods - analysis model, simulations and test-bed implementation. The existing research efforts done so far for EM of any class of sensor networks focus only upon simulations with application-based assumptions. The research finding observed from the qualitative literature review carried out for this research was that communication hardware also plays a vital role in overall energy consumption in the network. The amount of energy consumed during an active state of a node not only depends upon the metrics of radio but also on the drain efficiency of circuitries such as an inverter, Power Amplifier (PA) and so on. Therefore, energy modelling should be done based on all the parameters, which directly or indirectly cause higher energy consumption. In the work contributed by [20], an algorithm for maintaining perpetual operation of EH-WSN through maximisation of network throughput using reinforcement learning has been proposed. The other significant contributions by [21], [22], [23] and [24] also deal with energy optimization to achieve mitigation of dead nodes and higher energy consumption of WSN.The contribution by[25] deals with an energy efficient data reporting for navigation in position-free hybrid WSN and [26] present a comprehensive review on energy management schemes in EH-WSN.

3 ENERGY MANAGEMENT MODEL OF RF-ENABLED SENSOR NETWORKS

This sections briefs about the proposed methodology for the energy management and modelling of the RSN nodes followed by how it is solved using Semi Markov Decision Process (SMDP) and dynamic programming approach. This section also describes a fundamental explanation of these modelling methods to solve the issue of EM.

3.1 Markov Decision process

This subsection throws light into the basics of Markov Decision Process (MDP)and how can modelling be done based on MDP. MDP is defined as a controlled optimization process that results in providing a solution for the actions between nodes state transitions as defined by [16] and [27]. When an event/ action is triggered, the current state is transitioned to another state and a reward is presented to the state. This immediate reward presentation for each change in the states of the node ultimately leads to average reward being calculated for the entire model to solve the issue of energy consumption and arrive at favourable results. The MDP constitutes of the following elements i) Policy sets ii) An agent which decides and selects the action-based set of policies iii) Policy-based matrix for transition- which saves the probability of a particular transition iv) An award/ reward function- is the resultant function that provides a reward when the probability of a state's transition is successful. On the contrary, a transaction failure will result in either a discount function or penalty v) Objective function- this is the function which is vital for the performance comparison between the policy sets to solve the optimization issue. These MDPs can be applied to smaller systems and can be solved either using enumeration, DP or RL.

The proposed EM model is characterized on the basis of SMDP. Semi Markov Process (SMP) is defined as the stochastic process that dedicates a random amount of time (not unity) at each transition. Except for this functionality, SMP and MDP are almost the same. In other words, it can be said that the time taken for each of the transitions marks the difference between SMP and MDP. If this time factor is expressed as an exponentially distributed random variable, then the stochastic process is termed as Continues Time Markov process (CTMP). The notable difference between SMP and SMDP is that in SMP the system does not return or jump back to the same state, whereas for SMDP jumping back to the same state is possible. MDP, as described earlier, comprises of policies sets, reward function in the form of matrices, the objective function, policy-based matrix for state transition and the prime decision policy maker. All these together constitute the basic framework for solving the MDP. Accordingly, let us assume that $\alpha(i)$ refers to the policy α which determines the action chosen during i^{th} state and all such policies are deterministic in nature. The Probability Transition Matrix (PTM) is distinctive based upon each of the state's policy chosen and for each transition node, an immediate award (Reward function, (R_f) is assigned. After the completion of the transition, an average of all the R_f is assigned an average reward function, (Avg(R_f)), which utilizes the Probability Reward Matrix (PRM). Conclusively, the prime decision maker is either a network agent or processor. The transition time between the states is determined using Transition Time Matrix (TTM).

3.2 Dynamic Programming

This subsection explains and provides insights into the tool that is to be utilized for solving the markov process. Many research contributions have explored and utilized Dynamic Programming (DP) as one of the most helpful tools in solving the MDPs. DP figures out a complicated problem into many simple subproblems that can be determined and stored in a finite memory [28]. The previous versions of the solution that was intended for a particular problem can be utilized from this memory to determine the results of near future rather than starting it from the scratch. Many inter-related sub-issues can cause clarification on the major problem, which makes DP advantageous than its counterparts. It employs either a top-down or bottom-up approach where top-down solves the problem first followed by checking for the solution in the existing tabulated list, if a solution exists it updates, utilizes and stores back the value in the list, otherwise it solves the problem and stores it in the table. Bottom-up approach focuses on solving the multiple sub-issues and integrating each of the solutions to solve the major complex issue. The energy modelling based on MDP is solved using DP for this research in IoT context. For optimization control, DP employs the computation of a value function for each of the states. The feature of handling the problem complexities is played out better by DP than enumeration technique.

4 CONTROLLED OPTIMIZATION MECHANISM FOR ENERGY MANAGEMENT RSN MODEL (EMRSN)

This sections explains the process and concept of controlled optimization based on stochastic approach. It also provides insights into the computation of optimal values to gain higher lifetime and throughput. The proposed energy model has two different aspects for modelling the energy consumed by the RSN nodes. First one is the modelling during EH phase followed by ET state. Basically, any sensor node will have or operate in any one of its different operative states such as active, semi-active, idle, sleep, process or transmit/ receive. This research deals with modelling the RSN node in an energy efficient manner by switching between different states and by scheduling the duty cycling mechanisms for Tag Based Cooperative Solutions (TBCS). The prerogative nature of the sensor nodes is that during transmission and reception of data packets or during active state there is a lot of energy consumed. Whereas during idle or sleep state the relative energy expenditure is lesser. The semi-active state, on the other hand, works after a particular threshold level of the node's remaining energy is reached. The idle state is also more similar to the receiving state of the node, due to the fact that all the devices keep waiting to get the input. The fundamental system model of RSN during EH and ET is depicted in Figure 1 and 2 respectively. The two phases of the RSN node during and after deployment are EH and ET. Firstly, for energy harvesting systems, the harvesting happens at a low power rate during sleep mode and at a medium rate during idle mode. The transmission of packets is carried out during active state and the power saving / EH module is activated during semi-active state. The node's sense state is not made to function during EH, rather the process of sensing information is activated only during predefined time intervals followed by making the nodes to operate at lower power rate during harvesting, will also eventually improve the network lifetime. Secondly, for energy transfer systems there are totally 5 states in which energy modelling is carried out. The transmission and reception activities are done in the active state. In the semi-active mode, the node adapts to half of the energy ratings of the active mode. In sense state, the reader/ RF energy source is away from the node and therefore the sensed information is stored in the data buffer, by suspending the node from integrating with the RF source. In idle mode and sleep mode, the ET requests are queued up and in active mode, stochastic backscattering mechanism is carried out for ET, at

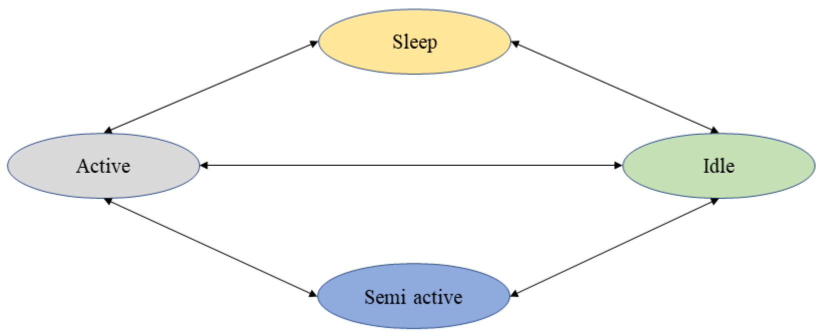

FIGURE 1
State transition diagram for Energy Harvesting RSN (EHRSN)

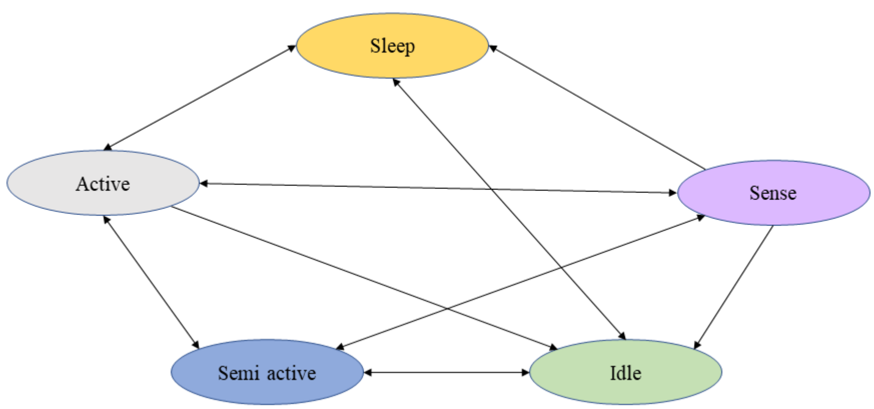

FIGURE 2
State transition diagram for Energy Transfer RSN (ETRSN)

States	Processor unit	Radio	T_x	R_x	Sensor unit	RFID unit	Total Harvesting Factor (THF)
Active (S0)	H_p	H_{Rad}	H_{T_x}	H_{R_x}	H_s	0	$H_p + H_{Rad} + H_{T_x} + H_{R_x} + H_s$
Semi active (S1)	$\frac{H_p}{2}$	H_{Rad}	H_{T_x}	H_{R_x}	$\frac{H_s}{2}$	0	$\frac{H_p}{2} + H_{Rad} + H_{T_x} + H_{R_x} + \frac{H_s}{2}$
Idle (S2)	0	H_{Rad}	H_{T_x}	H_{R_x}	0	$\frac{H_R}{2}$	$H_{Rad} + H_{T_x} + H_{R_x} + \frac{H_R}{2}$
Sleep (S3)	0	0	0	0	0	H_R	H_R

TABLE 1
Energy expenditure modes for EH

high power rate.

4.1 Elucidation of EM in RSN nodes

The elucidation of EM in RSN nodes has been briefed in this subsection. Fundamentally sensor nodes comprise of components such as a processor, sensing unit, radio unit, micro controller, RFID unit, and storage unit. Therefore, all of these components need the energy to power up their circuitries. This section further describes and determines the Total Harvesting Factor (THF) and Total Transfer Factor (TTF) for making each state of the node to run in an energy efficient manner. The Table 1 and 2 determines the energy modelling based upon the energy utilization factor for each corresponding state during EH and ET phases respectively. In tables 1 and 2 PU stands for processor unit, RU refers to the RFID unit and SU is the sensing unit. THF and TTF are the total harvesting factor and total transfer factor respectively. These factors determine the cumulative energy utilization factor during harvesting and energy transmission. The tabulation indeed depicts the various states of each process (EH & ET) along with the status of the circuits involved such as the status of the processor, radio, RFID, and sensing unit. Depending upon the type of process being carried upon, the utilization factor changes. For exam-

States	Processor unit	Radio	T_x	R_x	Sensor unit	RFID unit	Total Transfer Factor (TTF)
Active (S0)	T_P	T_{Rad}	T_{T_x}	T_{R_x}	T_s	T_R	$T_P + T_{Rad} + T_{T_x} + T_{R_x} + T_s + T_R$
Semi active (S1)	$\frac{T_P}{2}$	T_{Rad}	T_{T_x}	T_{R_x}	$\frac{T_s}{2}$	$\frac{T_R}{2}$	$\frac{T_P}{2} + T_{Rad} + T_{T_x} + T_{R_x} + \frac{T_s}{2} + \frac{T_R}{2}$
Sense (S2)	$\frac{T_P}{2}$	T_{Rad}	T_{T_x}	0	T_s	0	$\frac{T_P}{2} + T_{Rad} + T_{T_x} + T_s$
Idle (S3)	0	T_{Rad}	0	0	0	$\frac{T_R}{2}$	$T_{Rad} + \frac{T_R}{2}$
Sleep (S4)	0	0	0	0	0	0	0

TABLE 2
Energy expenditure modes for ET

ple, during active and semi-active states the nodes do not opt for harvesting process. The nodes get completely charged during sleep state and half charges during the idle state. We indefinitely assume that energy will be generated for sensors during the sleep state. For ET, on the other hand, there is no energy transmission during idle or sleep state. There is another state called sense state where RFID unit is completely disconnected and the operation of the sensor is to only sense the data and store it in its data buffer.

4.2 Policy Evaluation using Mathematical modelling

This subsection deals with policy evaluation using stochastic mathematical modelling for maximum optimality. The process that involves SMDP modelling for various states during EH and ET, is to evaluate the optimization policy. The SMDP framework is designed to solve the Markov decision problem. The seven basic elements of this framework are- States of the node, Policies, Actions, Transition probability functions, Transition reward functions, a decision maker and an objective function. The characteristics of each of the elements have been tabulated in Table 3. The decision-making component usually executes the action to be processed in each state of an SMDP. Each action taken for the transition from the state i to j is associated with PTM. The

States	In EH there are 4 states semi-active(S0), idle (S1), sleep (S2), active(S3). In ET there are 5 states Active (S0), semi- active(S1), Sense(S2), Idle (S3) and sleep (S4)
Action	In each state, there is an action (a, i, j) taken when a transition happens between the state Si to Sj
Transition probabilities	The following time take for a decision is a probability distribution function p(j)
Reward function	The state transitions are awarded with a reward function and it is computed based on Bellmans equation
Decision maker	This is responsible to select the control mechanism and is also termed as controller or agent
Policies	The policy is referred to as the control mechanism. A policy for a SMDP with n states is called n-tuple. Every element in this n-tuple determines the action to be opted during the current state of that element. If α is the policy, then for i^{th} element, α (i) refers to the action selected in i^{th} state.

TABLE 3
Characteristics of Nodes

transition in a Markov chain is associated with a reward function, which corresponds to the immediate cost implied for change in the state of the nodes. The control optimization problem consists of a performance metric called an objective function to compare the policies in terms of cost and reward factors. Average reward is termed as the expected reward function over an infinitely longer Markov processes calculated per unit time. The general average reward of a policy α can be determined by equation 1,

$$\mu_{\alpha=} \sum_{j \in s} \gamma_{\alpha}\ (j)\, p\,(j, \alpha\,(j)) \tag{1}$$

Where,$\gamma_{\alpha}\ (j)$ states the limiting probability distribution function when Markov process is run using policy α. S refers to the entire set of states in the Markov chain. $p\,(j, \alpha\,(j))$ relates to the immediate reward expected and earned during state j. This research utilizes the dynamic programming concept to solve the SMDP using controlled optimization technique. As stated earlier, the average reward policy is associated as a scalar quantity with every policy of a Markov process. Correspondingly, value function is utilized in the form of associative vector for each policy. This vector comprising of numerical values of the components can be solved by means of a linear set of equations, which is termed as Bellman equation. Therefore, in the context of average reward it is, determined as,

$$k_{\alpha}\,(j) = p\,(j, \alpha\,(j)) - \mu_{\alpha} + \sum_{i=1}^{|s|} r\,(j, \alpha\,(j)\,, i)\, k_{\alpha}(i), forall\ j \in s \tag{2}$$

$$k\,(j) = p\,(j, a, i) + \int \sum_{j \in s} \left[\int_0^x e^{-pt} \beta\,(j, a, i)\, r\,(j, a, i) \right] T dv \tag{3}$$

The above set of linear equations is equalized to the total number of elements in the set S, as | S|. Clearly, the two different DP- based methods for solving SMDP are either through policy iteration (PI) using Bellman equation for a policy or value iteration (VI) by utilizing the Bellman optimality equation. In equation 2 and 3, T is the total time distribution and the average reward is calculated using exponential distribution.$\alpha\,(j)$ refers to the action corresponding in state j for policy α and $r\,(j, \alpha\,(j)\,, i)$ determines the probability of one state transition for jumping from state j to i, using α policy which can be obtained from PTM. For the advantage of reducing the computational complexities, the Bellman Equation of Optimality (BEO) based upon VI algorithm using DP is employed in this research. This is also due to the fact that PI has many unknowns, leading to needless solving of many equations. The following equation presents the BEO,

$$k^*\left(j\right) = \ max_{a\in A(j)}[p\left(j,a,i\right) - \mu^* t\left(j,a,i\right) + \sum_{i=1}^{|s|} r\left(j,a,i\right) k^*\left(i\right)]), where\, j \in s \tag{4}$$

In the above equation 4, A(j) determines the entire set of actions permissible in state j. k* denotes the components of value function vector $\overrightarrow{K}^*$ which equalizes to the number of states in the SMDP. $p\left(j,a,i\right)$ refers to the immediate reward expected for selection of action a in state j to transition towards the next state i. $r\left(j,a,i\right)$ determines the PMT values for transitioning from state j to i on selection of action a and $\mu\,*$ denotes the average reward as stated in equation 1. According to SMDP the transition time between the state is non-exponential and deterministic. The reward function is calculated over T time distribution belonging to S, which specifies the instant of time when a transition should occur. t (j, a, i) represents the time taken to make a transition from state j to i when action a is taken. β(j,a,i) denotes the rate of reward. The values of p (j, a, i), r (j, a, i) and t (j, a, i) are computed and stored in matrix form termed as Reward Matrix for Transition (RMT), Probability Matrix for Transition (PMT) and Transition Time Matrix (TTM).

$$k^*\left(j\right) = \ max_{a\in A(j)} \left[p\left(j,a,i\right) - \mu^* t\left(j,a,i\right) + \ \rho \sum_{i=1}^{|s|} r\left(j,a,i\right) k\left(i\right) \right] where\, j \in s \tag{5}$$

Where ρ stands for discounted reward factor for negative awarding.

4.3 Optimal policy computation

This section expalins about how an optimized policy is calculated and computed to improve the node's lifetime following the identification of the energy expenditure details of a node in various states [28]. DP is utilized for solving the SMDP for the computation of optimized policy [29]. Random topologies of RSN have been extensively tested for time-based cooperative systems using temperature and humidity sensors on which the authors are currently working and is under research. The immediate reward for an RSN node is calculated using the following formula 6,

$$p\left(j,a,i\right) = \frac{THF/TTF\ \ (utilizationfactor)}{[(Energytime)] + \epsilon} forall\; i,j \in S \tag{6}$$

And where ϵ is the buffer for receiving/transmitting. The above equation is assigned to the transition of states by the node, based upon the energy utilization factor during EH and ET, explained in previous sub-sections.

4.4 Reward policy function for state transition

This subsection explains and briefs the computation of reward policy function for state transitions during the process of EH and ET. In this research, incentives in the form of rewards are awarded for favourable transitioned states. On the contrary, if the particular state transition does not provide energy optimization, negative reward or penalty will be imposed which is termed as a discounted reward. Therefore, efficient energy resource utilization will sum up for positive reward and vice-versa. On the transition from state j to i, immediate rewards are awarded under action a. For example, r (S_0, a) denotes immediate reward r in current state S_0, under action a. The following Table 4 and 5 shows the rewards awarded for each of the states in EH and ET processes.

States	**Rewards (R)**
Active (S0)	$2 \leq R \leq 4$
Semi active (S1)	$2 \leq R \leq 8$
Idle (S2)	$2 \leq R \leq 8$
Sleep (S3)	$4 \leq R \leq 16$

TABLE 4
Reward factor for EHRSN model

The illustration of assigning reward function for state transitions has been depicted in the form of the tuple in the following sub-section of this chapter using DP and Bellman's equation for optimality. Energy modelling based upon discounted reward factor is out of the scope of this research. The normalization of reward function is done on a scale of 0 to 1.0 by selecting 0.9 as the maximum value for the active state in both EH and ET. On the other hand, for more utilization of energy resources due to failed or uninitiated transaction, a higher penalty is imposed.

States	Rewards (R)
Active (S0)	$2 \leq R \leq 16$
Semi active (S1)	$2 \leq R \leq 8$
Sense (S2)	$2 \leq R \leq 8$
Idle(S3)	$2 \leq R \leq 4$
Sleep (S4)	$2 \leq R \leq 4$

TABLE 5
Reward factor for ETRSN model

4.5 Problem formulation and solution

In this sub-section, the bellmans equation is solved based on DP. Therefore, equation 2 is solved by computing RMT, PMT and TMT as depicted in Figure 3 and 4 respectively. For all the states of EH and ET processes, PMT, RMT, and TMT are derived and solved using DP based controlled optimization approach. The four scenarios for solving the stated SMDP are tabulated in Table 6 as follows,

Process	Energy/ Power Levels	Notation
EHRSN	Highly powered/ highly available	HP
EHRSN	Low powered/ low availability	LP
ETRSN	Highly powered/ highly available	HP
ETRSN	Low powered/ low availability	LP

TABLE 6
Four scenarios for solving SMDP

From Table 1 and 2, it can be evidently stated that all the state's operation and configurations are different for EH and ET. Since the aim of this research is to manage the energy levels, the process of EH is carried out during sleep state and idle state to maintain the balance between mitigation of higher en-

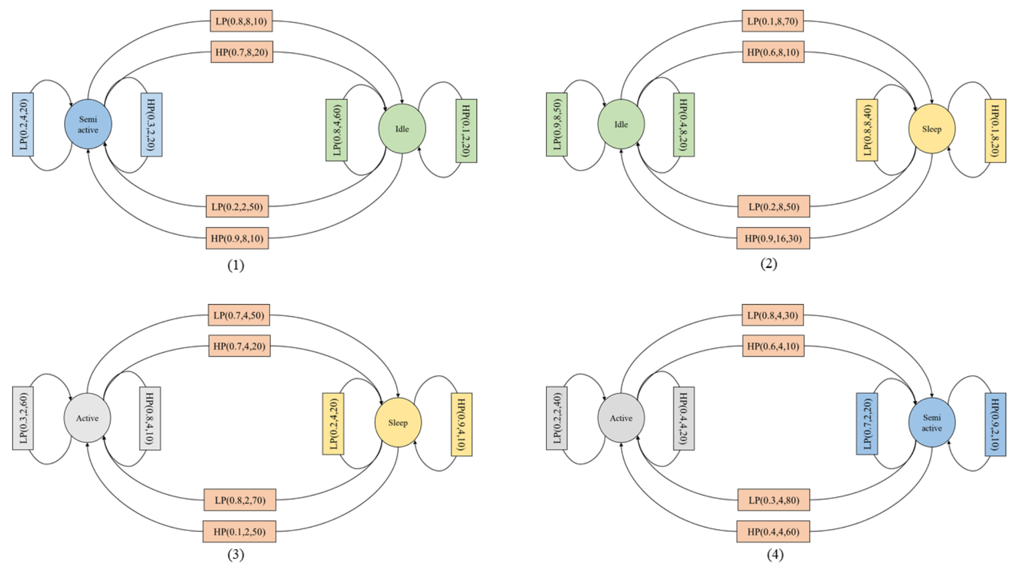

FIGURE 3
Illustrations of state transitions for EHRSN modeling

ergy consumption rate for T_xR_x operations and to prevent node failure/dead state of nodes. The nodes switch from idle to sleep state dynamically, when the network does not handle any tasks, or when there is no energy demand from the neighbouring nodes. This dynamic state switching is done based upon the time synchronization of nodes and the residual energy levels. The notation tabulated in Table 6 describes the energy levels high/low for EH and ET processes respectively. The illustrations depicted in Figures 3 and 4, shows the various state transitions between two states based on the energy levels (highly powered/low powered). These illustrations are further used to mathematically solve the SMDP process. H_p refers to highly available node, highly energized/highly powered mode stating that the node is highly active and does not opt for sleep state ultimately consuming more energy. L_p refers to the mode of having low energy levels/low powered indicating that the node's availability is also low that is at predefined intervals leading to the node being in a sleep state for harvesting energy. The other states active, sense and semi-active make the node to jump to the idle state when there is a need for lesser resource utilization. Therefore, each of the states opts for transitioning to an idle state for a considerably lesser number of times in order to avoid energy wastage during idle listening. Figure 1 and 2 depicts the STD of EH and ET processes for tag-based systems where the transition is represented by a notation of H_p and L_p to indicate the energy levels/availability followed by a tuple notation indicating the transition probability, immediate reward and time taken for the transition respectively as in the following equation 7,

$$H_{\mathrm{p}}/L_{\mathrm{p}} < P_{\mathrm{mt}}, R_{\mathrm{mt}}, T_{\mathrm{mt}} \ > \tag{7}$$

The option of SMDP makes the state transition possible along with the jumping back to its own state when compared to MDP. The transition between the different states of both EH and ET processes are computed mathematically using Bellman's equation and solved numerically based on DP. In this

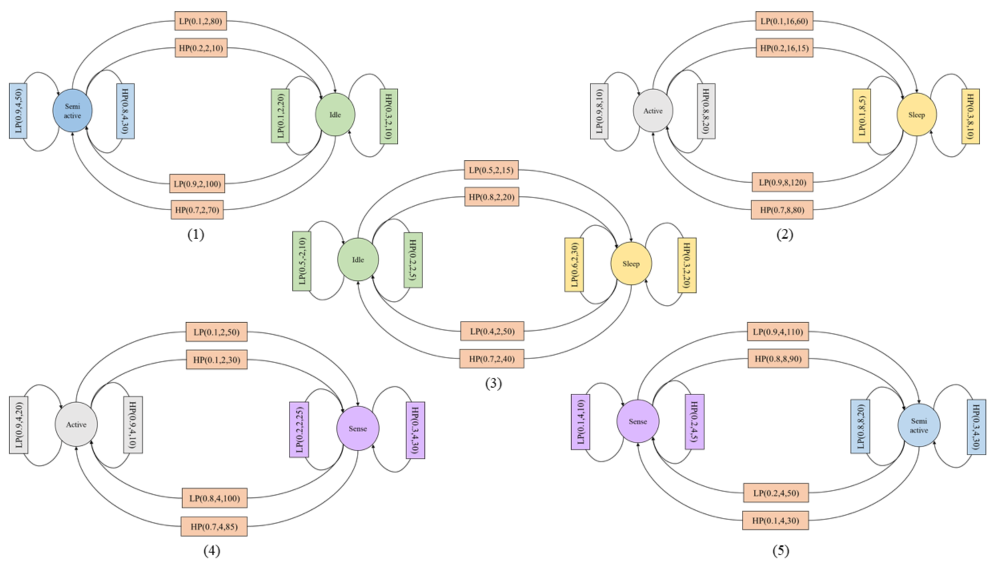

FIGURE 4
Illustrations of state transitions for ETRSN modeling

research, the optimal solution that can be achieved using consideration of the state's favourable transition is employed for effective resource allocation over a specific period of time. Hence, the policy evaluation algorithm is utilized to optimize the energy level and consumption during EH and ET processes of cooperative tag-based systems. For the purpose of policy improvement, the number of iterations is set to be, k and the number of states as S. The policy selection is made arbitrarily and for assumption basis, $O_{p(k)}$ is considered as the optimal policy achieved after computations. In the following equation 8, the variables t_k and μ_k are unknown, due to which either of the two should be replaced by 0. Since μ_k corresponds to the reward, it cannot be equalized to 0, therefore, the equation is solved by replacing t_k to 0. The transition time t (j, a, i) is also considered during policy computation and evaluation. $O_{(P(k)}$ is the selected policy with k number of iterations, and the new improved policy is chosen such that,

$$O_{P(K+1)}\left(J\right) \in argmax_{a \in A(j)} \left[p\left(j, a, i\right) - \ \mu^{k} + \sum_{j=1}^{|s|} r\left(j, a, i\right) t^{k}\left(j\right) \right] \tag{8}$$

If $O_{P(K+1)} = O_{(P(k)}$, then the chosen policy is said to be optimal and conclusive which stops the computation/evaluation of optimal policy. On the contrary, when there are no further improvisations with regards to the values of the policy's iteration, the computation furthermore, continues until the optimal policy is achieved. All the state transitions are considered for iterative policy evaluations for both the EH and ET processes, to compute the optimal policy and thereby achieve the most energy efficient solution.

5 NUMERICAL ANALYSIS AND RESULTS

The numerical analysis is carried out for both EH and ET process of TBCS by solving the SMDP process. The DP is employed for tabulating the values of transition probabilities ranging from 0.5 to 0.9 along with reward function of each of the transitions made from one state to another in the favour of harvesting and transferring energy followed by the time taken for each transition.

5.1 Numerical analysis of EH systems

This subsection explains the numerical analysis of the EH systems followed by tabulations of the optimal values obtained during the various state transitions for optimized energy harvesting in RSN. The equation 6, shows the

	A		B		C		D	
Low power	**Semi active to Idle**	**Reward**	**Idle to sleep**	**Reward**	**Active to sleep**	**Reward**	**Active to semi active**	**Reward**
$P_m = \infty$	-1.510	0.330	-2.610	0.200	-2.341	0.021	-1.710	0.431
P_m=0.5	6.441 0→0;1→1	0.129	6.576	0.242 0→1;1→1	9.717	0.210 0→1;1→1	2.761	0.079 0→1;1→0
P_m=0.6	1.815 0→1;1→1	0.169	4.00	0.400 0→1;1→1	1.205	0.222 0→1;1→1	3.964	0.072 0→1;1→1
P_m=0.7	3.852	0.380	0.112	0.340 0→0;1→1	4.240	0.234 0→1;1→0	4.632	0.207 0→1;1→0
P_m=0.8	0.489	0.479	3.640	0.632	3.000	0.280	2.972	0.318
P_m=0.9	1.886	0.434 0→1;1→1	4.597	0.524 0→0;1→1	6.1711	0.171 0→0;1→0	0.750	0.270 0→1;1→0
P_m=Actual probability	0.489	0.442	3.640	0.581 0→1;1→1	3.124	0.259 0→1;1→1	2.975	0.231 0→1;1→1
High power	**Semi active to Idle**	**Reward**	**Idle to sleep**	**Reward**	**Active to sleep**	**Reward**	**Active to semi active**	**Reward**
$P_m = \infty$	-1.2	0.16	-1.32	0.23	-1.28	0.291	-1.21	0.151
P_m=0.5	3.136	0.184 0→1;1→0	0.500	0.450 0→1;1→0	0.043	0.300 0→0;1→0	3.529	0.188 0→1;1→0
P_m=0.6	2.393 0→0;1→1	0.244	2.108	0.483 0→0;1→0	2.819	0.274 0→1;1→0	0.279	0.094 0→0;1→0
P_m=0.7	0.944	0.249	9.013	0.375 0→1;1→0	0.503	0.213 0→1;1→0	5.058	0.167 0→1;1→1
P_m=0.8	0.705	0.235 0→1;1→1	6.233	0.574 0→1;1→1	2.400	0.272 0→1;1→1	3.483	0.217 0→1;1→1
P_m=0.9	2.116	0.237 0→1;1→1	8.888	0.266 0→1;1→0	9.584	0.203 0→1;1→1	14.297	0.294 0→0;1→1

TABLE 7
Optimal values and average reward for EHRSN

notation assigned for preference of energy levels and numerical data analysis for the probability of successful transition, reward and time taken for the transition from current state to the next. For instance, in the first STD of Figure 3, the notation L_p (0.8,8,10) indicates that on the preference of low energy level mode, the transition probability for semi-active state to move to idle state is 0.8 for which higher rewards are given and since the transition happens from operational state to being idle, the time taken is much faster, 10 secs. Therefore, in low power mode, the preferable state for harvesting more efficiently is to transition to the idle state rather than staying in the semi-active state itself, for which L_p (0.2,4,20) holds true. In this method, the numerical analysis is done for all the states during EH and ET processes. The factors contributing to energy efficiency are given more attention to, during the analysis. The last element which is the transition time considered to be a unique case of SMDP is based upon the probability and frequency of switching between the states. The tabulations are shown in Table 7 and 8 depict the metrics related to the rewards and the corresponding probability values (P). The value of P varies from 0.5 to 0.9 followed by the actual probability for the system. The evaluation and computation include two states where reward metric indicates incentives to be awarded for transitioning between the states. 0 corresponds to initial/current state and 1 indicates the second state.

	A		**B**		**C**		**D**		**E**	
Low power	**Semi active to Idle**	**Reward**	**Idle to sleep**	**Reward**	**Active to sleep**	**Reward**	**Active to sense**	**Reward**	**Sense to semi-active**	**Reward**
$P_m = \infty$	-1.23	0.010	1.34	0.002	-1.114	0.014	1.76	0.128	-1.10	0.017
P_m=0.5	0.789	0.131 0→0;1→0	5.958	0.018 0→1;1→1	37.104	0.453 0→0;1→0	13.333	0.272 0→0;1→0	1.259	0.118 0→0;1→1
P_m=0.6	0.911	0.162 0→0;1→0	1.392	0.050 0→1;1→0	10.880	0.202 0→0;1→1	7.166	0.233 0→1;1→1	3.538	0.147 0→1;1→0
P_m=0.7	4.136	0.130 0→1;1→0	0.103	0.015 0→0;1→0	13.592	0.269 0→1;1→0	1.152	0.272 0→0;1→1	4.778	0.143 0→0;1→0
P_m=0.8	5.770	0.105 0→1;1→0	5.927	0.043 0→1;1→1	22.199	0.490 0→0;1→0	6.895	0.285 0→0;1→0	3.965	0.166 0→0;1→1
P_m=0.9	6.019	0.181	1.989	0.084 0→0;1→0	22.379	0.498 0→0;1→1	3.835	0.294 0→0;1→0	3.943	0.106 0→0;1→0
High power	**Semi active to Idle**	**Reward**	**Idle to sleep**	**Reward**	**Active to sleep**	**Reward**	**Active to sense**	**Reward**	**Sense to semi-active**	**Reward**
$P_m = \infty$	-1.12	0.021	-1.487	0.029	-1.298	0.126	-1.783	0.034	-1.875	0.110
P_m=0.5	2.480	0.131 0→0;1→0	2.857	0.114 0→0;1→0	23.320	0.294 0→0;1→0	16.437	0.338 0→0;1→0	1.394	0.017 0→0;1→1
P_m=0.6	6.809	0.218 0→0;1→0	4.370	0.059 0→0;1→0	15.891	0.277 0→0;1→0	39.024	0.199 0→0;1→0	2.574	0.067 0→1;1→1
P_m=0.7	4.147	0.147 0→0;1→0	0.262	0.041 0→1;1→0	9.060	0.199 0→1;1→0	17.749	0.246 0→0;1→1	3.509	0.103 0→0;1→0
P_m=0.8	8.636	0.233 0→1;1→0	1.129	0.159 0→0;1→0	37.263	0.526 0→0;1→1	10.066	0.397 0→1;1→0	2.000	0.428 0→0;1→0
P_m=0.9	6.727	0.219 0→0;1→0	7.870	0.149	20.951	0.353 0→0;1→0	15.190	0.366	4.078	0.370 0→1;1→1
P_m=*Actual probability*	6.727	0.228	2.019	0.153	20.951	0.383	23.481	0.420	27.136	0.413

TABLE 8
Optimal values and average reward for ETRSN

5.2 Numerical analysis for ET

This subsection explains the numerical analysis of the ET systems followed by tabulations of the optimal values obtained during the various state transitions for optimized energy transfer in RSN. The SMDP energy modelling provides the means for preferring the state and energy level that could favour the most for the efficient handling of EH and ET processes. The Bellman's equation is solved where $0 \rightarrow 0$ and $1 \rightarrow 0$ result denotes that irrespective of any probability chosen the state 0 is highly favourable. Accordingly, Figure 3 and 4 denote that, for EH, Idle and Sleep state are the most favourable states followed by Active state being the most favourable for ET process. The EHRSN and ETRSN framework are set in compliance with the total amount of energy that is harvested (THF) and the total amount of energy transferred (TTF). The basic assumption is that energy will be generated for sensors during the sleep state. All the corresponding units should be powered up completely according to the utilization factors THF and TTF. The Sleep state for ET process is negligible since all the units go to complete sleeping mode without any functioning or sensing operations. Table 7 and 8 show the combination of different rewards for various probabilities of state transitions, indicating that these configurations should be considered for real-time deployment of RSN devices. The average reward calculation is done for the values of Table 7 and 8 based on equation 6 and simulation based dynamic programming mechanism.

All the probability for state transition are considered from 0.5 to 0.9 and the reward function values are plotted for graphical analysis as depicted in Figure 5 and 6 respectively. It can be evidently concluded that when the probability of staying in a particular state is higher, the corresponding reward function is also more. The energy modelling is done based upon SMDP and is solved using dynamic programming approach. The state with probability P (∞) with probability (1,0) is also tested for an unreal case, where the optimal value tends to be in negative and the average reward comparatively smaller when compared to other limiting probabilities. In the P(actual) mode, the optimal state for idle, active and semi-active is sleep state due to the reason of energy harvesting being carried out during sleep and idle states. For ET, in the P (actual) mode, the most optimal state for idle, sleep and sense state are active and semi-active states respectively, due to the obvious reason, that ET using backscattering is carried out during the node's active state. For TBCS, Sleep state is the optimal one during EH since there is no any transmission/ reception activities leading to energy wastage or expenditure. For ET, in the actual mode the node is seldom modelled to go to sleep or idle state, rather it is kept in the active state for efficient transmission of energy through backscattering. It can also be seen from the numerical analysis that the sensor nodes work best and the average reward is higher when the transition probabilities are

EHRSN	**Idle**	**0.479** for low power 0.249 for high power with P_m= 0.8
	Sleep	**0.632** for low power 0.574 for high power with P_m= 0.8
	Active	**0.280** for low power 0.272 for high power with P_m = 0.8
	Semi-active	**0.318** for low power 0.294 for high power with P_m>= 0.8
ETRSN	**Idle**	0.181 for low power **0.233** for high power with P_m >=0.8
	Sleep	0.084 for low power **0.159** for high power with P_m >=0.8
	Active	0.498 for low power **0.526** for high power with P_m>= 0.8
	Sense	0.294 for low power **0.397** for high power with P_m>= 0.8
	Semi-active	0.166 for low power **0.428** for high power with P_m= 0.8

TABLE 9
Average reward for probabilities greater than 0.8

greater than 0.8 for both the processes as tabulated in Table 9.

6 EXPERIMENTAL RESULTS AND FINDINGS

This sections explains about the experimental results obtained after the modelling and also briefs the description of findings from the experimentation. The limiting probabilities determine the average reward for a particular state and also optimality to stay in a state or to make a transition towards next state. The tabulations and numerical analysis clearly state that for EH process, the probability of 0.8 or more works best at the lower power level. Moreover, for ET, the sensors tend to work more when the probability is either 0.8 or more at higher power levels. For example, if we assume that the probability of being in the active or semi-active consumes more energy and the probability to move to idle or sleep state is 0.9, then the sensors are made to harvest energy for a longer time in sleep mode or idle mode since there are no transmission or reception of data packets. According to Figures 5 and 6, two different motes are modelled according to the process of EH and ET through SMDP and they are solved using DP. Based upon the energy utilization factor during each of the process, immediate rewards were given. Thereafter, the average reward was computed using the DP approach and the energy utilized areas of both the process are analyzed numerically. An optimal policy is stated as that which maximizes the operation time of these motes during both EH and ET before going to complete dead state and getting depleted in their energy levels. The THF and TTF are calculated according to the process carried out by the sensor node.For ET process, the active state needs more energy to perform the activities of packet transmission and data processing, based upon the application. During EH process, the policy states that in sleep mode more energy can be harvested during non-sensing activities. However, the system prefers sleep state rather than idle state based upon the energy demands and utilization factors. However, during this period, there is a possibility of losing some information and queueing of data packets may again lead to overhead and energy consumption. This can amount to unfavourable modeling. Hence, the process of numerical analysis is being carried out in this section by considering such factors for TBCS and the actual normalized values are considered for calculation of average reward. In order to validate the said EM model, the RSN nodes are programmed as per the THF and TTF designed and formulated throughout the course of this paper. The real values of the temperature and humidity sensors adapted from their data sheets are approximated according to their respective utilization factors for suitable experimentation and validation. Besides, considering the inherent trade-off between throughput optimization, network longevity and energy optimiza-

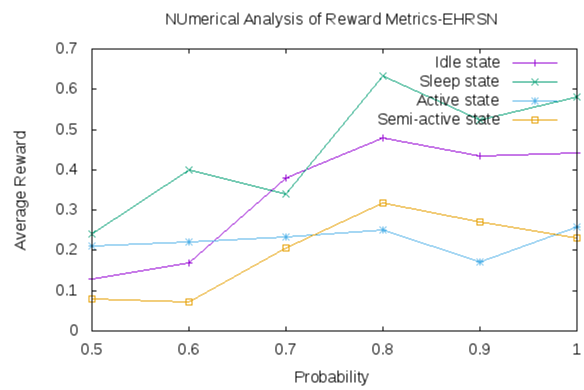

FIGURE 5
Average reward for EHRSN

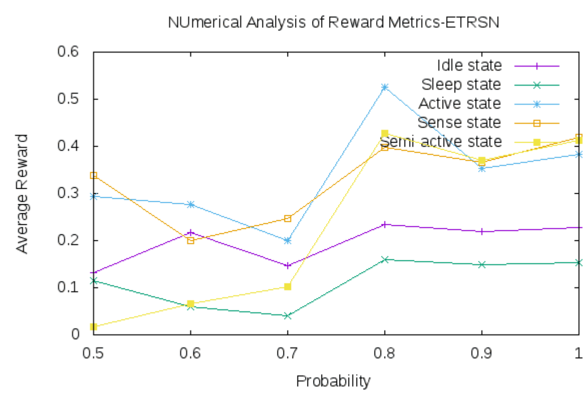

FIGURE 6
Average reward for ETRSN

tion when using the above described modelling aspects, the paper converts the trade-off to the average reward of being in a particular mode and accessing the EH and ET modules determined by residual energy levels and network

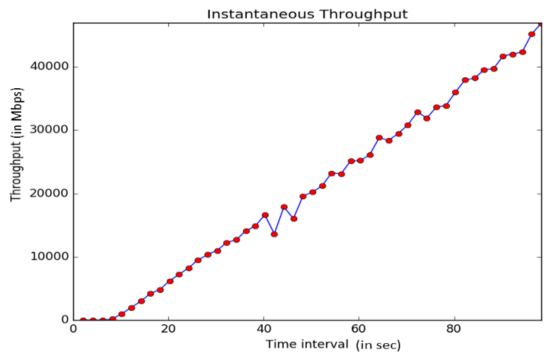

FIGURE 7
Instantaneous throughput after energy optimization

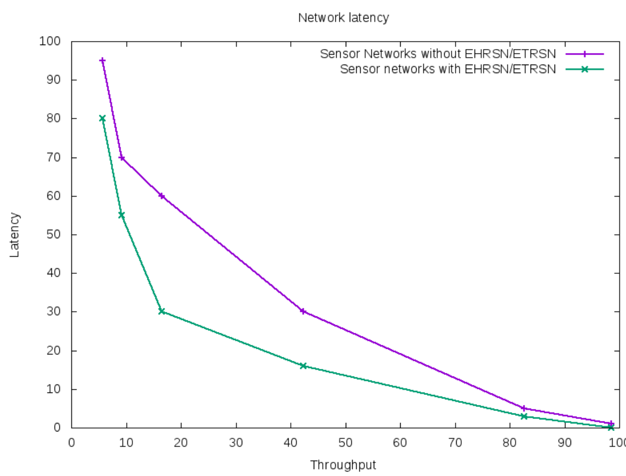

FIGURE 8
Network latency w.r.t throughput for sensor nodes with and without EMRSN modelling

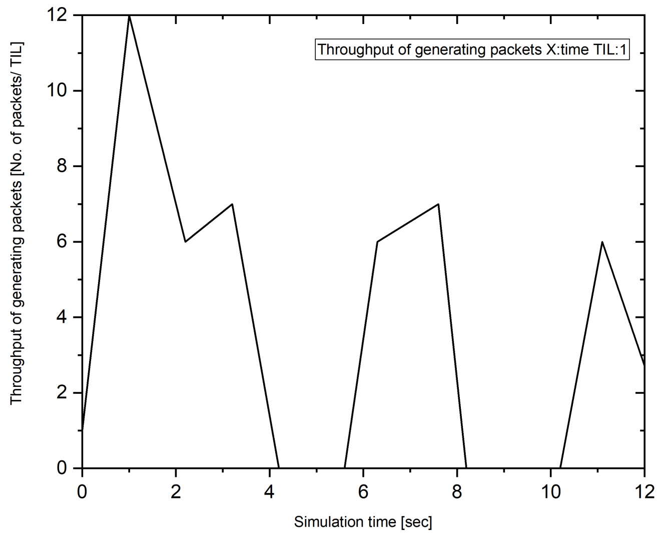

FIGURE 9
Throughput of generating packets during proposed optimization

performance. The proposed optimization validates that different application demands between various states and between throughput and lifetime can be adjusted accordingly by regulating the trade-off factor and coefficient of energy efficiency. Our findings are that when WSN is combined with RF, there is increasing instantaneous throughput as depicted in Figure 7 and it behaves differently with distinctive MAC interfaces. As shown in Figure 8, the integrated RSN (WSN+RFID) after implementation via simulation is tested for network throughput with and without the modelling mechanism of EMRSN. For this purpose, the integrated WSN + RFID, WSN, and RFID are compared in terms of energy consumption and residual energy to analyze the energy consumption factor of the integrated network. Here, the term ZigBee is used during some instances, instead of WSN since the MAC for both is the same (802.15.4 MAC) and the term tag here refers to the RSN nodes because tags are embedded together with the ZigBee sensor nodes in network simulations.

On the other hand, it is obvious that when both are combined they consume more energy due to sensing, processing, and reading capabilities. The proposed model is focused and implemented to manage this energy expenditure when both RFID + Zigbee are combined. The graphical representation show that the throughput is increased on combining WSN with RFID when compared with WSN alone which means that WSN being entirely dependent on energy storage devices like batteries and capacitors are bound to have decreased throughput levels rather than WSN with ambient energy harvesting of RF signals. Moreover, the network lifetime is increased by nearly 80%

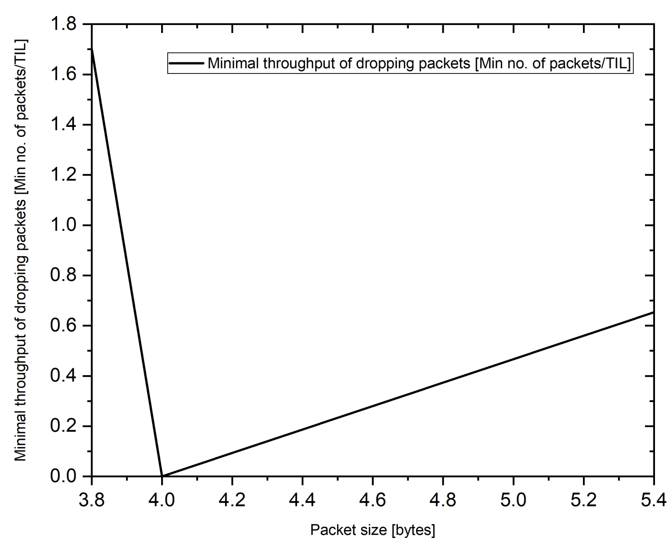

FIGURE 10
Packet size Vs minimum throughput upon optimized harvesting

rather than WSN without RF harvesting. It was also found that the MAC and the physical layer parameters along with the upper layers can play a significant role in identifying the energy profile of a sensor node. Ultimately, the rate at which harvesting takes place depends upon the size of the node being recharged or the size of the packet being generated, dropped or transferred as

depicted in Figure 9, 10 respectively. As shown in Figure 9, the throughput of generating packets by the RSN network with respect to simulation time follows a random distribution and provides maximum network efficiency when there is process of EH carried out by the nodes. During the active state, the nodes provide maximum throughput after EH and during ET process also, the event-triggered scheduling is carried out in order to ensure decreased network delay and packet drop. The throughput increases whenever there are packet generation and transmission, which implies that the network latency is lower indicating towards higher network efficiency. The relationship between packet size and throughput of sending packets as depicted in Figure 10 clearly implies that the total time required to read the data from the RSN nodes is directly proportional to the memory size of the node.Therefore, the size B bits of the data and energy buffer of a particular node implies comparatively smaller time to recharge the node through harvesting or backscattering process. These results clearly signify that throughput of generating and sending packets during the process of EH and ET are random and exponential in nature. The optimized harvesting process also signifies that the packet size generated is directly proportional to the throughput.
Hence, in a nutshell, the research articulated in this paper is threefold-firstly, the model is described for both EH and ET process, followed by evaluation of optimal policy using mathematical modeling and solving the Bellmans equation based on DP. Secondly, the reward functions are then calculated based upon the probability of a states transition and the time taken for the transition. Thirdly, the average reward is further calculated based on problem formulation and numerical analysis of SMDP for TBCS. Finally, the sensors are then programmed accordingly to achieve maximum throughput, network lifetime and energy efficiency.

7 CONCLUSION AND FUTURE WORK

In this paper, the development of proposed EM model for RSN has been presented and evaluated, to solve the challenge of higher energy consumption and limited network lifetime. The RSN nodes are characterized as SMDP and are solved using a controlled optimization mechanism with simulation-based dynamic programming. Firstly, the model is described for both EH and ET process, followed by evaluation of optimal policy using mathematical modelling and the Bellmans equation based on DP is solved. The reward functions are calculated based upon the probability of a states transition and the time taken for the transition. The average reward is further calculated

based on problem formation and numerical analysis of SMDP for TBCS. The aim of achieving optimal energy consumption and mitigating the complete dead state of nodes during EH and ET process has been achieved. The authors are currently developing and implementing algorithms for EH and ET of RSN and also are validating it with the hardware model with longer distance between the RF source and sensor node.
The future work of this research can be carried out to involving artificial neural network based detection for identification of decreased energy levels and employment of energy harvesting as described for WSNs in [30]. Similarly constraint based forwarding for multihop broadcasts when RFID and WSN are integrated can also be focussed upon to overcome the drawbacks of higher energy expenditure in sensor networks. The other interesting future directives for RSNs would be energy efficient precise localization and energy aware localised QoS routing as described by researchers in [31]. Another quintessential factor to be focussed upon when cross-layer technologies are integrated on an IoT platform is the security issue and addressing it using energy harvesting and transfer as discussed in [32].

8 ACKNOWLEDGEMENT

The authors would like to thank the Sustainable Science Cluster, University of Malaya (Grand Challenge Grant UM.0000007/HRU.GC.SS GC002B15SUS) for their support.
This project is also supported by RU Grant - Faculty Programme, GPF009D-2018.

REFERENCES

[1] Gaddafi Abdul-Salaam, Abdul Hanan Abdullah, and Mohammad Hossein Anisi. (2017). Energy-efficient data reporting for navigation in position-free hybrid wireless sensor networks. *IEEE Sensors Journal*, 17(7):2289–2297.

[2] Gaddafi Abdul-Salaam, Abdul Hanan Abdullah, Mohammad Hossein Anisi, Abdullah Gani, and Abdulhameed Alelaiwi. (2016). A comparative analysis of energy conservation approaches in hybrid wireless sensor networks data collection protocols. *Telecommunication Systems*, 61(1):159–179.

[3] Nabil Ali Alrajeh, Shafiullah Khan, Jaime Lloret, and Jonathan Loo. (2013). Secure routing protocol using cross-layer design and energy harvesting in wireless sensor networks. *International Journal of Distributed Sensor Networks*, 9(1):374796.

[4] Nabil Ali Alrajeh, Shafiullah Khan, Jaime Lloret Mauri, and Jonathan Loo. (2014). Artificial neural network based detection of energy exhaustion attacks in wireless sensor networks capable of energy harvesting. *Ad Hoc & Sensor Wireless Networks*, 22(1-2):109–133.

[5] Giuseppe Anastasi, Marco Conti, Mario Di Francesco, and Andrea Passarella. (2009). Energy conservation in wireless sensor networks: A survey. *Ad hoc networks*, 7(3):537–568.

[6] Shaik Shabana Anjum and Rafidah Md Noor. (6-7 September, 2018). Towards energy-efficient rf-enabled sensor networks in internet of things. In *Conference Proceedings IntelliSys*, pages 1–6. London, United Kingdom, IEEE.

[7] Shaik Shabana Anjum, Rafidah Md Noor, Ismail Ahmedy, Mohammad Hossein Anisi, and Norazlina Khamis. (29-30 November, 2017). Energy management techniques for rf-enabled sensor networks based on internet of things. In *International Conference on Computational Science and Technology*, pages 53–63. KualaLumpur, Malaysia, Springer.

[8] Shaik Shabana Anjum, Rafidah Md Noor, Mohammed Hossien Anisi, Ismail Bin Ahmedy, Fazidah Othman, Muhammad Alam, and Muhammad Khurran Khan. (2019). Energy management in rfid-sensor networks: Taxonomy and challenges. *IEEE Internet of Things Journal*, 6(1):250–266.

[9] Aliyu Aliyu Babayo, Mohammad Hossein Anisi, and Ihsan Ali. (2017). A review on energy management schemes in energy harvesting wireless sensor networks. *Renewable and Sustainable Energy Reviews*, 76:1176–1184.

[10] June-Ho Bang, Young-Jong Cho, and Kyungran Kang. (2018). An energy-efficient data delivery scheme exploiting network coding and maximum node-disjoint paths in wireless sensor networks. *Adhoc & Sensor Wireless Networks*, 42:87–106.

[11] Michael Buettner, Ben Greenstein, and David Wetherall. (March 30- April 01, 2011). Dewdrop: an energy-aware runtime for computational rfid. In *Proc. USENIX NSDI*, pages 197–210. Boston, MA.

[12] C-F Chiasserini and Michele Garetto. (2006). An analytical model for wireless sensor networks with sleeping nodes. *IEEE Transactions on Mobile Computing*, 5(12):1706–1718.

[13] C-F Chiasserini and Ramesh R Rao. (2001). Energy efficient battery management. *IEEE journal on selected areas in communications*, 19(7):1235–1245.

[14] Dina S Deif and Yasser Gadallah. (2017). An ant colony optimization approach for the deployment of reliable wireless sensor networks. *IEEE Access*, 5:10744–10756.

[15] Denis Dondi, Alessandro Bertacchini, Davide Brunelli, Luca Larcher, and Luca Benini. (2008). Modeling and optimization of a solar energy harvester system for self-powered wireless sensor networks. *IEEE Transactions on industrial electronics*, 55(7):2759–2766.

[16] Abhijit Gosavi. (2003). Simulation-based optimization. *parametric optimization techniques and reinforcement learning*, 55(2).

[17] Fabio Iannello, Osvaldo Simeone, and Umberto Spagnolini. (May 23-27, 2010). Energy management policies for passive rfid sensors with rf-energy harvesting. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–6. Cape Town, South Africa, IEEE.

[18] Mian Ahmad Jan, Priyadarsi Nanda, Xiangjian He, and Ren Ping Liu. (9 December, 2014). Pasccc: Priority-based application-specific congestion control clustering protocol. *Computer Networks*, 74(PB):92–102.

[19] Pouya Kamalinejad, Chinmaya Mahapatra, Zhengguo Sheng, Shahriar Mirabbasi, Victor CM Leung, and Yong Liang Guan. (2015). Wireless energy harvesting for the internet of things. *IEEE Communications Magazine*, 53(6):102–108.

[20] Jaime Lloret, Miguel Garcia, Angel Catala, and Joel JPC Rodrigues. (2016). A group-based wireless body sensors network using energy harvesting for soccer team monitoring. *International Journal of Sensor Networks*, 21(4):208–225.

[21] Xiao Lu, Ping Wang, Dusit Niyato, Dong In Kim, and Zhu Han. (2015). Wireless networks with rf energy harvesting: A contemporary survey. *IEEE Communications Surveys & Tutorials*, 17(2):757–789.

[22] Renita Machado, Nirwan Ansari, Guiling Wang, and Sirin Tekinay. (2010). Adaptive density control in heterogeneous wireless sensor networks with and without power management. *IET communications*, 4(7):758–767.

[23] Muhammad Mostafa Monowar. (2017). An energy-aware multi-constrained localized qos routing for industrial wireless sensor networks. *Adhoc & Sensor Wireless Networks*, 36(1-4):21–50.

[24] Dina M Omar and Ahmed M Khedr. (2018). Erplbc-cs: Energy efficient routing protocol for load balanced clustering in wireless sensor networks. *Adhoc & Sensor Wireless Networks*, 42(1-2):145–169.

[25] Karishma Singh, Karan Singh, Ahmed Aziz, *et al.* (19 June 2018). Congestion control in wireless sensor networks by hybrid multi-objective optimization algorithm. *Computer Networks*, 138:90–107.

[26] Maninder Singh and Paramveer Gill. (15-16 December, 2017). Wireless sensor networks with effective algorithm for long life and energy optimization. In *Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2017 International Conference on*, pages 32–34. Mysuru, India, IEEE.

[27] Rudd JM Vullers, Rob Van Schaijk, Hubregt J Visser, Julien Penders, and Chris Van Hoof. (2010). Energy harvesting for autonomous wireless sensor networks. *IEEE Solid-State Circuits Magazine*, 2(2):29–38.

[28] Yin Wu and Kun Yang. (30 April-1 May, 2018). Cooperative reinforcement learning based throughput optimization in energy harvesting wireless sensor networks. In *Wireless and Optical Communication Conference (WOCC), 2018 27th*, pages 1–6. Hualien, Taiwan, IEEE.

[29] Wei Xiang, Ning Wang, and Yuan Zhou. (2016). An energy-efficient routing algorithm for software-defined wireless sensor networks. *IEEE Sensors Journal*, 16(20):7393–7400.

[30] Feng Yang, Lin Du, Weigen Chen, Jian Li, Youyuan Wang, and Disheng Wang. (2017). Hybrid energy harvesting for condition monitoring sensors in power grids. *Energy*, 118:435–445.

[31] Pallavi Yarde, Sumit Srivastava, and Kumkum Garg. (4-6 December, 2017). A modified energy efficient protocol for optimization of dead nodes and energy consumption in wireless sensor networks. In *Sensing Technology (ICST), 2017 Eleventh International Conference on*, pages 1–6. Sydney, NSW, Australia, IEEE.

[32] Amir Ehsani Zonouz, Liudong Xing, Vinod M Vokkarane, and Yan Lindsay Sun. (2016). Hybrid wireless sensor networks: a reliability, cost and energy-aware approach. *IET Wireless Sensor Systems*, 6(2):42–48.