1. Introduction

Notably, many studies are now focusing on green communications for wireless sensor networks (WSNs). WSNs usually rely on a fixed battery to supply energy for data transmissions, and network lifetime is mainly dominated by the battery capacity. Several techniques are investigated to maximize and prolong the lifetime of battery-powered nodes [1–3] such as design new routing path, a mathematical model, and a novel algorithm. However, the proposed techniques may be inconvenient and costly, and may produce impractical results. Consequently, the lifetime of a battery-powered node is still limited. Therefore, energy harvesting (EH) has recently attracted significant attention in many applications such as military usage, air defense, civilian technology for the environment, and visual sensor networks used for surveillance cameras [4–6]. Each application has different implementation considerations, such as the environment, the application’s design objectives, cost, hardware, and system constraints. For a general introduction to WSNs, we refer the readers to the survey literature in [7]. Another aspect in applications is that the required signal is detected then transmitted through WSNs using limited energy resources. Further, resolving energy supply problems is very significant in wireless networks using hybrid energy to provide a perpetual and sufficient energy supply.

To date, there have been many studies focusing on circuit, hardware, and manufacturing for intelligent solar energy systems [8] to enlarge the strength of energy harvesting (EH) utilization. Additionally, various policies and models have been considered to increase the efficiency of EH...
utilization in WSNs [9–12]. In [9], a minimum-increment matched marginal method was proposed to maximize solar harvested energy utilization. In [10], the authors attempted to maximize harvested energy utilization by using an expectation-maximization algorithm in parallel with a Markov decision process (MDP). In addition, the energy prediction model for multi-source EH was presented in [11] to expect future energy availability based on the past energy observations. Additionally, the prediction model used real-life solar and wind traces to improve the performance of the prediction accuracy. In [12], a novel policy for energy-efficient transmission in harvesting sensors was adopted to expect the sum of the utility values only for transmitted important messages. This policy used threshold function with MDP and depends on the battery level. Based on numerical experiments, the stochastic approximation algorithm was proposed to find the optimal policy that is more efficient and has a faster convergence speed than the Q-learning algorithm.

Various energy management methods were studied in [13–15] by considering different objectives such as the estimation and prediction of EH based on a full-information or partial-information knowledge of EH. In [13], a switching algorithm for rechargeable tree-based WSN was proposed using MDP. In particular, the proposed algorithm alternatively switches from one sensor node to another based on solar energy levels to guarantee the network sustainability and energy efficiency for sensor nodes. Moreover, a real-time sensor traffic pattern was used to analyze the energy consumption of sensor nodes. In [14], the source estimation for EH of sensor nodes was studied for both deterministic and stochastic modes. For deterministic mode, the amount of EH is known prior to transmission for the deterministic model. Besides, the optimal offline power allocation was considered to minimize the average mean-square-error (MSE) over the estimation periods. For stochastic mode, the online power allocation with Lyapunov optimization was considered to minimize the MSE over the estimation periods, where only the current amount of EH is known. In [15], event capture in rechargeable sensor networks was investigated for two policies. The first policy considers a full-information model to find the optimal greedy policy based on dynamic control theory. The second policy considers the partial-information model which gets information when the event occurs only at an active mode of the sensor. With an efficient heuristic clustering policy, the second policy uses partially observable MDP to obtain the optimal solution.

On the other hand, hybrid energy utilization techniques were adopted in [16–18] to guarantee the reliability of the WSN connection. A dynamic programming method was designed in [16] to find the optimal and suboptimal solutions for point-to-point communications with hybrid energy. A new class of hybrid networks is investigated in [17] to evaluate the performance of hybrid energy utilization. This class evaluates the hybrid energy performance based on the outage duration and the transmission duration criteria, where EH, conventional nodes, and channel fading statistics affect the aforementioned criteria. Under another perspective of hybrid energy, two cluster head technologies are proposed [18], which are powered by solar energy and wireless charging to provide high power density without causing any risk to health. To maximize hybrid energy utilization, a polynomial-time scheduling algorithm and a linear-time algorithm are considered to make a trade-off between solar energy and wireless charging for the cluster head.

In [19–22], Quality of Service (QoS) research issues were investigated that focused on different objectives. In order to determine the optimal location of the energy-harvesting sensor nodes (EHSNs) and the battery-powered sensor nodes (BPSNs), an effective location selection protocol was proposed in [19]. Moreover, the proposed routing protocol was studied for different QoS requirements on maximizing the average end-to-end path reliability, cost of communication paths for both EHSNs and BPSNs, and energy consumption for applications running on hybrid WSNs. In [20], the distortion minimization problem for EH in WSN was investigated to optimize the harvested energy utilization. To solve this problem, the MSE of distortion minimization was formulated using a low-complexity sleep–wake scheduling and power control algorithm. In addition, both limited orthogonal channels and the prior knowledge of the channel power gain were proposed for each sensor node. Additionally, a discrete-time Markov chain was applied in [21] to analyze the average
packet dropping probability (PDP) under two specific automatic repeat retransmission protocols without optimizing the transmission policy. On the other side, the trade-off between the energy consumption and the packet error probability were investigated in [22] for body sensor networks (BSNs). In order to maximize the quality of coverage for BSNs, an MDP is formulated by using EH, taking into account both the number of dead slots and the average number of not-reported messages.

To the best of our knowledge, there have been no works considering the tradeoff between the minimization of the constant energy source utilization and the PDP with the usage of a hybrid energy source through the MDP method in WSN. In this article, hybrid energy utilization consisting of a constant energy source and solar harvested energy is achieved by considering an MDP design framework. Due to the unpredictability of solar harvested energy, a constant energy source is used to ensure reliable data transmissions. In fact, more constant energy source consumption is required for reducing the PDP if the solar harvested energy is insufficient. Our aim is to find the transmission policy that optimizes the trade-off between the minimization of constant energy source consumption and PDP.

The rest of this paper is organized as follows. Problem formulation and Markov decision process are clarified in Section 2, including a case study. Simulation results are presented in Section 3 via computer simulations. Finally, conclusions are drawn in Section 4.

2. Problem Formulation and Markov Decision Process

2.1. System Model

The system uses hybrid energy to transmit data packets in a WSN, as shown in Figure 1. Here, the data is periodically reported from each sensor to the sink node. Moreover, two buffers are considered, in which one represents a battery for solar harvested energy and the other indicates a buffer for data packets. In order to achieve reliable communications, a constant energy source is required if the harvested energy in the battery is insufficient. Otherwise, the event of dropping a packet occurs given that the packet buffer is full or the solar harvested energy in the battery is not enough to transmit any packet.

We study the optimal transmission policy for hybrid energy utilization concerning the battery and packet buffer states, actions, transition probabilities, and cost value for each action. The hybrid energy utilization is formulated as an MDP design framework for simple wireless communication between one sink and one WSN node in order to find the optimal transmission policy.

The MDP model can be described as \((S, A, P, R)\). Let \(S\) be the state space which is a composite space of the packet state \(Q = \{0, \ldots, N_Q - 1\}\) and battery state \(B = \{0, \ldots, N_B - 1\}\); i.e., \(S = Q \times B\), where \(\times\) denotes the Cartesian product. At the state \(S = (i, j)\), it means that there are \(i\) packets and \(j\) available solar energy units in the packet buffer and the battery, respectively. \(A = A_C \times A_H\) is the composite action for transmitting packets using hybrid energy, including the action for transmitting...
packets using the constant energy source $A_c = \{0, 1, \ldots, N_A\}$ and the action for transmitting packets using solar harvested energy $A_H = \{0, 1, \ldots, N_A\}$, where $N_A$ is the maximum allowable action index. We assume that $P$ is the state transition probability for different actions, and $R$ is the reward function to harmonize the cost value of constant energy source usage and the PDP for all actions.

2.2. Action

We assume that each packet transmission consumes one energy unit. That is, if an action $(a_C, a_H) \in A$ is taken, the total number of transmitted packets is equal to $a_C + a_H$. At the state $S = (i, j)$, the affordable actions can be classed into the following types:

1. Constant energy source only when $a_H = 0$ and $a_C \in \{1, \ldots, N_A\}$.
2. Solar harvested energy only when $a_H \in \{1, \ldots, \min\{j, N_A\}\}$ and $a_C = 0$.
3. Hybrid energy when $a_H \in \{1, \ldots, \min\{j, N_A\}\}$ and $a_C \in \{1, \ldots, N_A\}$.
4. No transmission when $a_H = 0$ and $a_C = 0$.

2.3. System States Transition Probabilities

Since the events of packet arrival and solar harvested energy arrival are independent of each other, the transition probability from the current state $(S_Q, S_B) = (i, j)$ to the next state $(S_Q, S_B) = (i', j')$ with respect to the action $A = (a_C, a_H)$ can be represented by the following equation:

$$
P_A((S_Q, S_B) = (i', j')|(S_Q, S_B) = (i, j)) = P_A((S_Q = i')|(S_B = i)) \cdot P_A((S_B = j')|(S_Q = j))
$$

(1)

A two-state Markov chain model is considered here for describing the packet arrival and the solar energy arrival [10]. If an action $(a_C, a_H)$ is taken, the battery and packet state transition probabilities $P_A((S_Q = i')|(S_Q = i))$ and $P_A((S_B = j')|(S_B = j))$ can be specified as follows:

$$
P_A((S_Q = i')|(S_Q = i)) = \begin{cases} 
\eta_P, & i' = i - (a_C + a_H) + 1; \\
1 - \eta_P, & i' = i - (a_C + a_H); \\
0, & \text{Otherwise.}
\end{cases}
$$

(2)

$$
P_A((S_B = j')|(S_B = j)) = \begin{cases} 
\eta_S, & j' = j - a_H + 1; \\
1 - \eta_S, & j' = j - a_H; \\
0, & \text{Otherwise.}
\end{cases}
$$

(3)

where $\eta_S$ and $\eta_P$ are the probabilities of one solar energy unit arrival and one packet arrival, respectively.

2.4. Cost Function

The cost function is defined as the weighted cost values for both the constant energy source consumption $a_C$ and the PDP. Specifically, the cost function at the state $S = (S_Q, S_B) = (i, j)$ with respect to the action $A = (a_C, a_H)$ is defined as

$$
C_A(S) = \begin{cases} 
(1 - \alpha)\eta_P, & i = N_Q - 1, a_C = a_H = 0; \\
\alpha \cdot a_C, & \text{Otherwise.}
\end{cases}
$$

(4)

The parameter $\alpha$ is the weighting factor, ranging between zero and one, to make a performance tradeoff between the constant energy source consumption and the PDP. Besides, the PDP occurs only when the packet buffer is full, i.e., $S_Q = N_Q - 1$, and a packet arrives, and meanwhile, the composite action equals to 0 (i.e., no packet is transmitted from the packet buffer). In this case, the cost of
dropping a packet is given by \((1 - \alpha) \eta_p\). On the contrary, the cost of using a constant energy source is \(\alpha \cdot a_C\) when packets in the buffer are transmitted. Accordingly, the PDP is emphasized if \(\alpha\) approaches zero. On the other hand, if \(\alpha\) is set approaching to one, it reflects the importance of saving the constant energy source.

### 2.5. Optimal Transmission Policy

Our goal is to find the optimal transmission policy \(\pi^*(S)\), which specifies the optimal action at the state \(S\), to make a tradeoff between the minimization of the expected cost values of dropping packets and using constant energy source. The optimal policy satisfies Bellman’s equation as shown in [4]:

\[
V_{\pi^*}(S) = \min_{A \in A} \left( C_A(S) + \gamma \sum_{S' \in S} P_A(S'|S) V_{\pi^*}(S') \right), S \in S
\]

(5)

where \(\gamma\) is a discount factor, and \(V_{\pi^*}(S)\) represents the expected cost value with respect to the optimal policy \(\pi^*(S)\). The optimal policy that satisfies the Bellman’s equation can be found via a well-known value iteration algorithm [23]. At the \(i\)th iteration, we have

\[
V_{i+1}^{A}(S) = C_A(S) + \gamma \sum_{S' \in S} P_A(S'|S) V_{i}(S'), S \in S, A \in A
\]

(6)

\[
V_{i+1}(S) = \min_{A \in A} \{V_{i+1}^{A}(S)\}, S \in S
\]

(7)

The value of \(V_0(S)\) is initialized as zero for all states \(S \in S\). The above procedures in (6) and (7) are repeated for several iterations until a stopping criterion is met; i.e., \(|V_{i+1}(S) - V_i(S)| \leq \epsilon\), where \(\epsilon\) is a preset threshold.

### 2.6. Case Study

This subsection presents a case study of the proposed scheme. In the MDP formulation, it is possible to transmit or receive more than one packet at a time, but in this case (as shown in Figure 2), it is only allowed to transmit or to receive one packet at most, where the MDP model is \(N \times N\) states for battery and packet buffers, actions, and the state transition probabilities. In this model, the maximum value of arrival or departure rates for both packet and battery unit is one per cycle. The actions are classified as follows: the black dotted line indicates no packet transmission action; the red dashed line indicates the transmission action of one packet using one constant energy source unit; the green dashed line indicates the transmission action of one packet using one battery unit from the solar harvested energy, and the violet dashed line indicates no transmission action with the dropped packets.
3. Simulation Results

In this section, the performance of the proposed hybrid energy utilization scheme is evaluated by computer simulation based on a value iteration algorithm, as shown in Algorithm 1. Two performance metrics—(1) a constant energy utilization ratio and (2) a packet dropping probability—are taken into consideration. Based on [10], we assume the discount factor ($\gamma = 0.99$), stopping criterion = $10^{-6}$, and the cost value for the transmission of one packet or dropping one packet is one energy unit individually. For simplicity, we assume that the packet arrival to the wireless sensor buffer can be zero or one. In each time slot, the proposed MDP model uses both battery units and the constant energy source to transmit a number of packets such as zero, one, two, three, and so forth.

**Algorithm 1: Value Iteration**

**Input:** Initial Cost Values $C_A(S)$

**Output:** Optimal Expected Cost Value $V_{\pi'}(S)$

Begin the loop;
Calculate expected cost values for first state $V_1(S)$
Calculate expected cost values for second state $V_2(S)$

while $|V_{i+1}(S) - V_i(S)| \geq \epsilon$ do

Calculate expected cost values for next state $V_{i+1}(S)$

repeat

$\quad i \leftarrow i + 1$;

until

$|V_{i+1}(S) - V_i(S)| < \epsilon$

end the loop

$V_{\pi'}(S) = V_{i+1}(S)$

---

**Figure 2.** Markov chain for the Markov decision process ($N_Q = N, N_B = N$).
Figure 3 presents a flowchart for the value iteration algorithm, which finds the optimal transmission policy based on the proposed MDP model. In a conventional method (the performance calculations of hybrid energy utilization without using MDP), the expected value is as the summation of cost value for each action that will be taken over the same loop, similar to the value iteration algorithm’s loop for the proposed MDP method. The scenario which does not use MDP has no transmission policy, and the action will be taken randomly for each time slot. Besides, it does not have the discount factor and stopping criterion to calculate the expected value. According to the configuration setting, two performance metrics (i.e., power consumption ratio and PDP) are compared with/without using MDP in two different simulation environments. In the first simulation environment, the model sizes of the battery and packet buffer are two by two, three by three, or four by four. Additionally, the solar energy unit arrival is assumed to be only zero or one (Case 1), which fits the aforementioned model sizes as depicted in Figures 4 and 5.

In Figure 4, the difference of the power consumption ratio between the MDP scenario and the scenario without MDP can reach up to 40%. In both scenarios, the power consumption ratio between the constant energy source and total hybrid power decreases as the model size increases. The reason behind this is that the larger battery or packet buffers accommodate more solar harvested energy for transmitting packets. Consequently, more packets can be transmitted in a timely fashion, avoiding packet dropping. However, in the without-MDP scenario, the average power consumption ratio can be fixed because the weighting factor $\alpha$ does not affect the constant energy source usage. On the other hand, Figure 5 shows that the PDP decreases for both scenarios as the model size increases. In the MDP scenario, this led to more packets being received and transmitted at the right time based on solar harvested energy. In addition, the percentage of PDP will be fixed when $\alpha$ is larger than or equal to 0.5 for both scenarios. Needless to say, this percentage does not change, since the cost of PDP

![Flowchart for the value iteration algorithm.](image-url)
will be always less than the cost of a constant energy source used in this case. On the other side, $\alpha$ will be negligible without using MDP and the action will be dropping the packet.

![Figure 4](image)

**Figure 4.** Power consumption ratio for different values of $\alpha$. MDP: Markov decision process.

![Figure 5](image)

**Figure 5.** Packet dropping probability (PDP) for different values of $\alpha$.

In the second simulation environment, the same model size of the battery and packet buffer is used, which is four by four, as shown in Figures 6 and 7. Additionally, the performance metrics with different energy quanta (discrete values) of the solar energy unit arrival are used. This creates various cases of solar energy unit arrival as follows:

Case 1: Arrival units could be zero or one.
Case 2: Arrival units could be zero, one, or two.
Case 3: Arrival units could be zero, one, two, or three.

The aforementioned figures show that the performance metrics improved when the energy quanta probability of solar harvested energy arrival increased for all cases. In Figure 6, the MDP scenario outperformed the non-MDP scenario up to 45% in terms of the power consumption ratio difference. Additionally, the percentage of power consumption ratio is decreased in Cases 2 and 3 rather than in Case 1. In particular, the battery utilization based on solar harvested energy is improved significantly in
the MDP scenario. In addition, $\alpha$ reflected on the MDP scenario results, since each action is determined by the cost function of the minimum cost value. In contrast, the average power consumption ratio is fixed when the MDP scenario is not considered. In both scenarios, Figure 7 shows that the PDP decreases when the probability of solar harvested energy arrival increases—especially in Cases 2 and 3, where the battery could accumulate more solar energy to transmit packets without dropping a packet. Moreover, the percentage of PDP will be high and fixed for the MDP scenario when $\alpha$ is larger than or equal to 0.5. Based on the value of $\alpha$, the cost value of PDP will be higher than the cost of constant energy source for all cases. In the scenario without MDP, $\alpha$ will be negligible, and packets are dropped for all cases.

![Power Consumption Ratio](image)

**Figure 6.** Power consumption ratio for different values of $\alpha$.

![PDP](image)

**Figure 7.** PDP for different values of $\alpha$.

Simulation results show that the design framework can efficiently balance the constant energy source usage and the PDP by adjusting the weighting factor $\alpha$, based on the proposed MDP model. When the value of the weighting factor $\alpha$ increases, the power consumption ratio decreases, since the cost of the constant energy source usage is high. On the contrary, the PDP increases, since the cost of the PDP is less important. In addition, when the value of $\alpha$ is larger than or equal to 0.5, the power consumption ratio is zero and the PDP remains constant. The reason for the tendency in figures lies in that the cost of the PDP is less important than the cost of the constant energy source usage when...
\(\alpha \geq 0.5\) according to (4). Additionally, hybrid energy utilization will be increased when the model sizes or the energy quanta probability of solar harvested energy arrival increase, as emphasized in previous figures.

4. Conclusions

This article focuses on efficient hybrid energy utilization for WSNs. Considering the unpredictability in solar harvested energy sources, the system model shows that the hybrid energy utilization can be optimized based on the proposed MDP model with the cost function, the constant energy source consumption and PDP. With a finite packet buffer and a finite battery size, the MDP design framework includes the states space, the actions set, the state transition probabilities, and the cost function to tradeoff between the minimization of the expected two cost values. Based on Bellman's equation, the optimal strategy is found for all states and actions. Simulation results confirm that the MDP model can effectively achieve the optimal solution for balancing the minimization of the constant energy source consumption and the PDP.

Acknowledgments: This work was supported by the Ministry of Science and Technology, MOST 106-2221-E-216-002 and MOST 106-2221-E-009-050, Taiwan.

Author Contributions: Mohammad Tala’t and Kai-Ten Feng arranged the literature, performed the simulation, analyzed the results and wrote the paper; Chih-Men Yu and Meng-Lin Ku suggested the problem and cooperated in designing the system model also helped to formulate the problem formulation and the proposed MDP model.

Conflicts of Interest: The authors declare no conflict of interest.

References


