October 2019

SWITCH STRATEGY FOR SMART UTILITY NETWORK ORTHOGONAL FREQUENCY-DIVISION MULTIPLEXING PHY BASED ON REINFORCEMENT LEARNING IN LOW-POWER AND LOSSY NETWORKS

Ling Wei

Wenjia Wu

Chuanwei Li

Leo Yu

SWITCH STRATEGY FOR SMART UTILITY NETWORK ORTHOGONAL FREQUENCY-DIVISION MULTIPLEXING PHY BASED ON REINFORCEMENT LEARNING IN LOW-POWER AND LOSSY NETWORKS

AUTHORS:
Ling Wei
Wenjia Wu
Chuanwei Li
Leo Yu

ABSTRACT

Techniques are described herein for a switch strategy based on actor-critic reinforcement learning. This depends on not only the Received Signal Strength Indicator (RSSI) and packet loss rate but also on the distribution of error segments in a packet. The states, actions, and values for the actor-critic model may be defined to fit the switch strategy. The mechanism of dynamical adjustment may depend on the different PHY formats.

DETAILED DESCRIPTION

The IEEE 802.15.4g protocol is applied in connected grid mesh networks (CG-MESH). This protocol supports multiple Smart Utility Network (SUN) Orthogonal Frequency-Division Multiplexing (OFDM) PHY formats to adapt the data rate on the basis of the wireless Channel State Information (CSI). In current CG-MESH systems, the switch strategy among these different formats depends on the Received Signal Strength Indicator (RSSI) and packet loss rate. However, packet collision may occur due to the "hidden node" problem. Packet collision may be misjudged due to channel fading in current switch strategies. Thus, the node may incorrectly decide to switch the SUN OFDM PHY to the lower data rate format, thereby degrading network performance and causing disruption in the system.

Accordingly, a switch strategy is provided herein to avoid misjudgment between packet collision and channel fading. The switch strategy depends not only on the RSSI and packet loss rate, but also on the distribution of error segments in a packet. Furthermore, actor-critic reinforcement learning is applied to determine the switch strategy and improve performance in Low-power and Lossy Networks (LLNs).

There are several SUN OFDM PHY formats in CG-MESH systems. When a wireless channel suffers from great attenuation or noise interference, which results in higher packet loss rate, the relevant nodes should switch to a lower data rate format to guarantee valid communication service. Conversely, nodes should improve the data rate when the channel improves. Therefore, the CSI may be accurately measured while packets are received.

Due to the "hidden node" problem, packet collision may occur, increasing the packet loss rate. But the relevant nodes should not switch the PHY format because doing so cannot resolve packet collision. Thus, the system may determine the true reason for increasing packet loss rate: packet collision or channel error.

Figure 1 below illustrates an example typical packet structure.

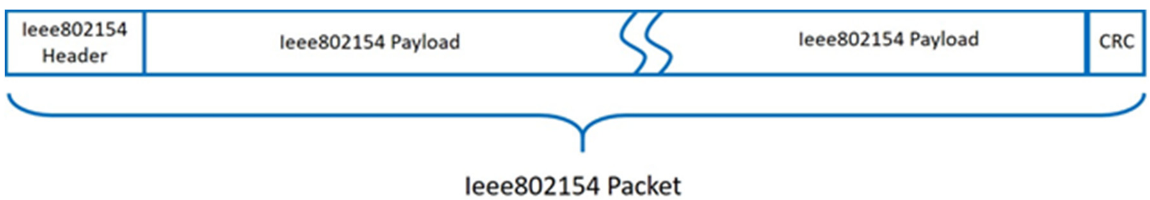


Figure 1

RSSI and packet loss rate alone is not enough to distinguish the true reason for packet loss, because the results of packet collision and channel fading are nearly identical. Accordingly, a multi - Cyclic Redundancy Check (CRC) segments packet structure is provided to offer more details regarding packet loss. Figure 2 below illustrates an example multi-CRC segments packet structure.

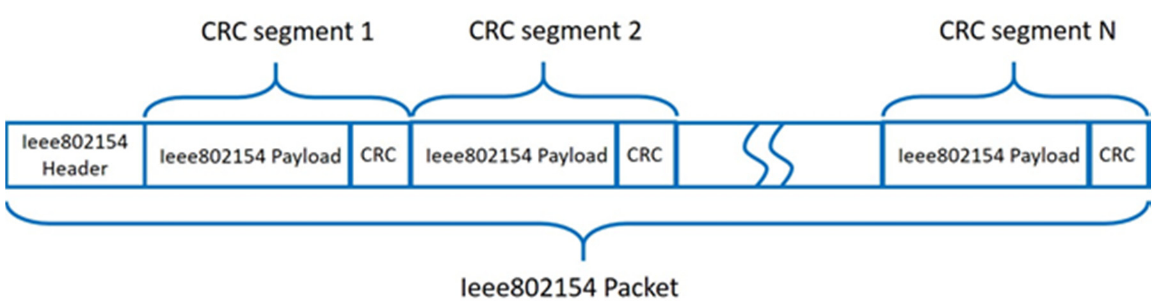


Figure 2

After receiving a packet, the CRC value in each segment may be calculated and the error distribution may be exported. This may allow determination of the cause of packet loss based on the differing distributions of packet collision and channel error. These techniques significantly improve the accuracy of such determination.

A statistical approach of the numbers of collision packets and channel error packets is provided herein. The statistical rules may be designed based on the minimum and maximum numbers of successive erroneous CRC segments ("num_min" and "num_max," respectively). As illustrated in Figure 3 below, the num_max for the collision packet scenario should be larger than the num_max for the channel fading scenario.

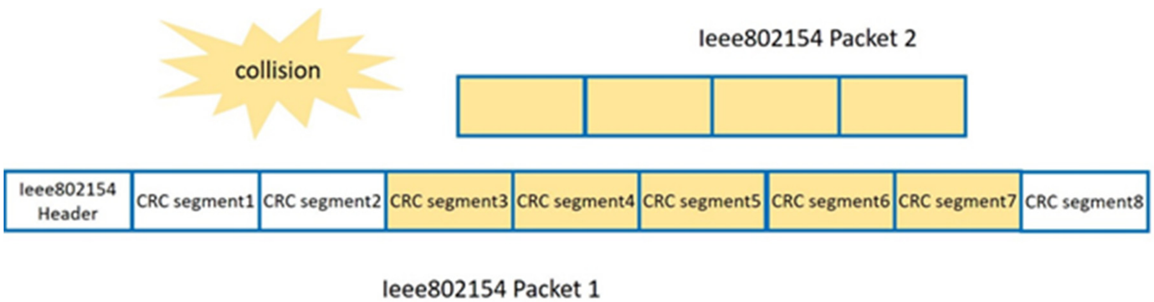


Figure 3

Meanwhile, as illustrated in Figure 4 below, the num_min for the channel fading scenario should be smaller than the num_min for the collision packet scenario.

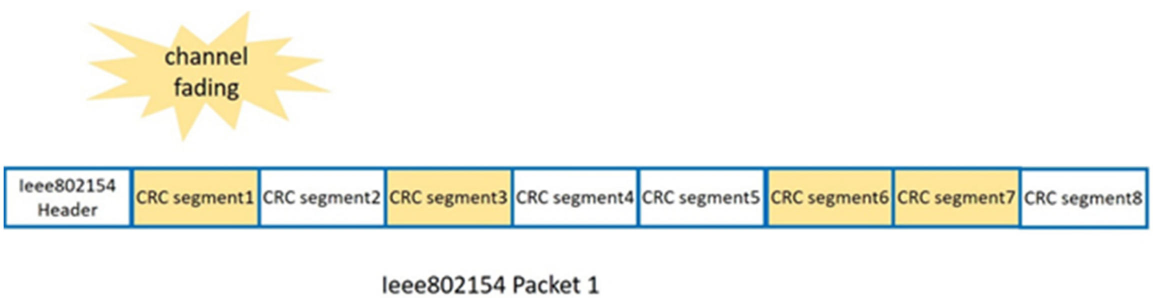


Figure 4

Example rules are provided as follows:

If (num_min <= 1 && count >= 4)

num_channel _error ++;

else if (num_max >=4)

num_collision_packet ++;

end

Statistics for multiple packets may also be obtained over a given period of time. In general, different rules may be designed according to the conditions specific to a particular system. Also/alternatively, the rules may be dynamically adapted in response to a changing environment. The number of successive erroneous CRC segments may be useful in many such situations.

Also described herein is an actor-critic reinforcement learning model to distinguish the reason for packet loss rate and make the switch decision based thereon. For instance, after obtaining the statistical results of num_channel_error and num_collision_packet, a reinforcement learning model may be designed to decrease erroneous judgment. In one example, an actor-critic model may be applied. The actor-critic model includes two networks: the actor-network, which inputs the states and outputs the actions, and the critic network, which inputs the states and actions and outputs the value.

Figure 5 below illustrates an example flowchart for an actor-critic reinforcement learning model. First, the states (used by the actor-network) may be generated. This may be (or be derived from) the statistical results discussed above. The states may be fit into the actor-critic reinforcement learning model to obtain the results (actions from the actor-network), which is the switch operation, and evaluate the rewards (values from critic network), which is the data rate increase. The policy may be improved depending on the rewards (e.g., increasing the gradient and modifying the model parameters).

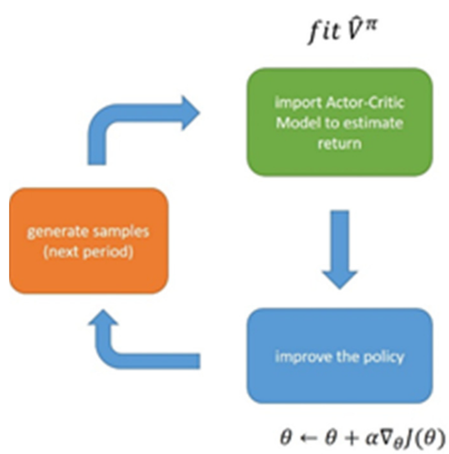
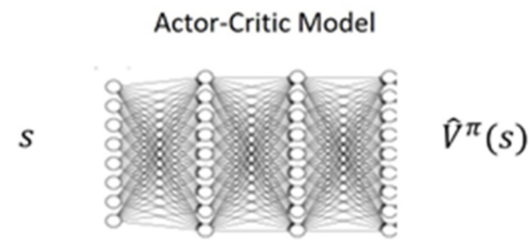


Figure 5

The performance of the actor-critic reinforcement learning model may improve with experience. Parameters may be automatically modified each cycle to improve performance in order to obtain more rewards. Figure 6 below illustrates an example schematic chart of an actor-critic reinforcement learning model.



$$\hat{V}^{\pi}(s_t) = \sum_{t'=t}^{T} E_{\pi\theta}[r(s_{t+1}, a_{t+1})|s_t]$$

Figure 6

The switch strategy may also be dynamically adjusted according to the subcarrier modulation mode of OFDM formats. Different OFDM PHY formats have disparate error-tolerant capabilities due to the character of their constellation diagrams. For Modulation and Coding Scheme (MCS) 5 and MCS 6, Quadrature Amplitude Modulation 16 (16 QAM) is used in subcarriers. 16 QAM supports a high data rate but low error tolerance. By contrast, MCS 4, MCS 3, and MSC 2 apply Quadrature Phase Shift Keying (QPSK) modulation. Thus, the distance between two constellation points increases, permitting the receiver to correctly demodulate the signal in spite of packet collisions.

Adjusting the statistical approach to adapt to the changes may improve the probability of misjudgment. When the system applies MCS 5 and MCS 6, the difference between channel error and collision may be large enough to distinguish MCS 5 and MCS 6 using a simple reinforcement learning model. However, when the MCS 3 and MSC2 are applied, a more sophisticated reinforcement learning model may be used and more judgment conditions may be considered. Dynamic adjustment may improve the throughput and robustness of the system.

In summary, techniques are described herein for a switch strategy based on actor-critic reinforcement learning. This depends on not only the RSSI and packet loss rate but also on the distribution of error segments in a packet. The states, actions, and values for the actor-critic model may be defined to fit the switch strategy. The mechanism of dynamical adjustment may depend on the different PHY formats.