Enabling intelligent traffic flow management in Wireless LAN's using Markov Decision Process tools

Sonali Aggarwal, Shrey Gupta, sonali9@stanford.edu, shreyg@stanford.edu
Under the guidance of **Professor Andrew Ng**
12-11-2009

1 Introduction

Current resource allocation methods in wireless network settings are ad-hoc and fail to exploit the rich diversity of the network stack at all levels. We want to deploy machine learning algorithms for assigning resources at the access points in Wireless LANs in a real-world wireless setup, to bring about significant improvements in throughput of the network. The new IEEE standard 802.11n Multiple-Antenna-Equipped Wireless LAN standard allows for numerous options at both the PHY and MAC layer of a transmitting wireless node, where a particular selection of options result in system performance maximization for a given state of the wireless channel, network, background traffic and user application. Our goal is to use the Markov model developed for the .11n standard and deploy machine learning tools to select the option at the transmitting node, from the available choices within the realm of the standard, that maximizes system performance, within the time limit dictated by the rate of change of the state-variables. We first deployed Markov decision processes (MDP) to choose the optimal policy for obtaining the maximize throughput. However, MDPs explore the entire state space to find the optimal policy (which is a slow process and since our network conditions keep changing frequently, markov model takes a long time to re-learn the state transition probabilities). We deployed online algorithms which are much faster. Our preliminary analytical results combined with MATLAB simulations have verified that our approach outperforms existing approaches.

2 Motivation

Wireless LANs are popping up everywhere to provide Internet access for the ever-more ubiquitous laptop-toting customer. They are easy to install and they give users the mobility to move around within a local coverage area and still be connected to the network. Plans are in place to cover entire cities with WLANs. The WLAN industry, as it is now, with deployments only in limited places, is over 2 billion with over 100 million unique users. With a massive widespread deployment expected to hit the consumers in a few years, the numbers are expected to grow by an order of magnitude or more.

IEEE 802.11 is a set of standards carrying out wireless local area network (WLAN) computer communication. IEEE 802.11n-2009 is an amendment to the IEEE 802.11-2007 wireless networking standard to improve network throughput over previous standards. Behind most 802.11n enhancements lies the ability to receive and/or transmit simultaneously through multiple antennas. The 802.11n protocol uses MIMO strategy for communication. Multiple Input Multiple Output (MIMO) communication is well-known to boost the wireless spectral efficiency through spatial multiplexing. At the physical (PHY) layer, advanced signal processing and modulation techniques have been added to exploit multiple antennas and wider channels.

The limiting factor in the performance of a WLAN today is resources at the access point. Current resource allocation methods are ad-hoc and fail to exploit the rich diversity in wireless network settings at all levels of the network stack. Maximizing the system capacity, by selecting the best policy in the WLAN (to increase the data-rate) is the motivation behind our project. We developed and deployed Machine learning algorithms for the resource allocation problem that seek to have a direct impact on WLAN system performance. The problem of finding the optimal control policy (amongst the various options available) to maximize throughput is cast into a Markov Decision Problem (MDP). Our preliminary analytical results combined with MATLAB simulations have verified that our approach outperforms existing approaches, but the next step is to try it on a real testbed.

The 802.11n Multiple-Antenna-Equipped Wireless LAN standard allows for numerous options at both the PHY and MAC layer of a transmitting wireless node, where any set of options results in system performance maximization for a given state of the wireless channel, network, background traffic and user application. For a typical WLAN deployment, the aforementioned state-variables change dynamically and at a fast rate. Our goal is to use the Markov model developed for the .11n standard [Bia00] and device a Markov Decision Process to select the option at the transmitting node, from the available choices within the realm of the standard, that maximizes system performance, within the time limit dictated by the rate of change of the state-variables. System performance in terms of throughput will be the criterion for performance.

Based on the work by Giuseppe Bianchi [Bia00], which derived a Markov model for state-variable change at both the MAC and PHY layers, we do reliable state-estimation using characteristics of the previously received packets (e.g. signal strength, number of retransmissions, packet loss rate, etc.) on top of any relevant header information provided by the standard. State-estimation for some of the variables can also be done via explicit feedback from the receiving node, allowed by the standard, furthermore the cost of such explicit feedback will be included in throughput calculation.

State and action variables used in our model

For the purpose of this class project, we plan to include the following state variables in our model

- **Packet Success Rate**: Packet success rate is a good indicator of channel strength, and is readily available via Ack's received from the receiving node, or via Ack Timeouts, whichever may be the case. We need to choose the optimal policy which gives us the maximum Packet Success Rate i.e. the maximum throughput.

The following actions constitute the action variable set in our MDP formulation

At the MAC layer:

- **MAC Layer Packet Frame Length**: The frame length of the MAC packet affects the packet reception probability. Longer frames transport more user-data for the same header length, however requires good channel quality for a longer duration of time. The standard provides for 10 different options for MAC frame length.

- **User Selection**: In a typical WLAN deployment, the transmitting node may be sending data to multiple receiving nodes. Selecting the right user plays an important role in overall system performance.

At the PHY layer, the actions are

- **Modulation and Coding Scheme (MCS)**: The 802.11n standard defines Modulation and Coding Scheme (MCS) a simple integer assigned to every permutation of modulation, coding rate, guard interval, channel width, and number of spatial streams. The 802.11n standard provides for 8 different options for the MCS. Identifying the MCS values supported in 802.11n devices is a good way to determine the set of data rates that can actually be utilized by users WLAN. If the state of the channel is good, MCS with a higher index ensures a higher data rate.

3 Implementation

3.1 The standard approach - MDP

The problem of finding the optimal action given the state of the wireless network can be essentially posed as a Markov Decision process problem. We divided the problem of finding optimal policies for all combinations of the state and individual actions into sub-problems . To start with, we have assumed a stationary model for simplicity which considers the wireless channel to be stationary.

We attacked the problem of finding the best policy by deploying the standard approach of repeatedly gaining experience and learning state transition function P in the given scenario and then converging to the optimal policy. It is **difficult to quantify the gains** achieved as at present the selection is done on an ad-hoc basis by the settings of the network controller and those controlling individual devices in the network. The stability of the approach assures us that we always select the action corresponding to the maximum throughput.

3.2 Alternate Approach - Online Learning

Though the MDP approach explained is a robust approach to deal with our problem formulation, we observed that in certain conditions it failed to converge and select optimal actions. This can be attributed to the conditions where the wireless channel is continuously changing and the experience accumulated by the MDP is not enough to allow it to converge to an optimal policy. Thus we essentially encounter a **exploration problem**, which marks a trade-off between learning time and performance of a MDP - the more time we spend in exploring the state-action space the better results we get.In wireless networks, the channel is inherently dynamic due to a number of factors like fading, interference among others which are not under the purview of the network itself. This motivates us to look for a approach that does not require a stationary characteristic from the network/channel.

Since we always aim to maximize the throughput by selecting the best action in minimal time, we decided to use an algorithm which can continuously give us a good action even while its learning. So we followed the 'online learning' approach to find the optimal policy. We adapt the 'online learning' to our problem scenario by considering an ordered pair of actions

(a_1, a_2) and the corresponding change in throughput (ΔTh). The change in throughput is used as a guide to select the next action. Value for the learning rate μ was selected according to the scale on which throughput was measured (and the sampling time of throughput), so as to have only scalable changes in the action. We use the fact that in best of the conditions the throughput is directly proportional to the the MAC layer packet length and the MCS level.

We do not directly use the current ΔTh value to determine our next action. We average the value of ΔTh observed over various transitions from a_1 to a_2 and use this ΔTh_{avg} to determine our new action. This approach can be translated as the following algorithm.

3.2.1 Online Learning Algorithm

Algorithm 1 Online Learning Algorithm

procedure SELECT-NEXT-ACTION($a_1, a_2, \Delta Th$) ▷ Action pair (a_1, a_2)
    if $a_2 > a_1$ **then**
        flag $= 1$
    else
        flag $= -1$
    end if
    Using the ΔTh evaluate the value ΔTh_{avg} for action pair a_1, a_2
    $a_{new} = a_2 + flag * \mu * \Delta Th_{avg}$
    Normalize a_{new} to limit the value within the action space
end procedure

4 Simulation

We successfully deployed the above algorithm on the Simulink model. To better visualize the performance of the algorithm we carried out simulations for the marginal cases in which the network is stationary. Figure 1 refers to one such case in which we have quasi-stationary state with fixed MAC layer collision probability = 0.5. Figure 2 lists the average throughputs observed for each of the possible actions of selecting the MAC Layer packet frame length.

Subsequent to a successful run of the algorithm on the Simulink Model, We are in the process of deploying the same algorithm on a real Test Bed and plan to submit the results for publication.

5 Model

The simulation were carried out on a Simulink model. The complete model and related code can be found at http://www.stanford.edu/~shreyg/CS229/

6 Collaborator

Rajiv Agarwal (rajivag@stanford.edu), PhD. Candidate (under Prof. John M. Cioffi), Department of Electrical Engineering, Stanford University.

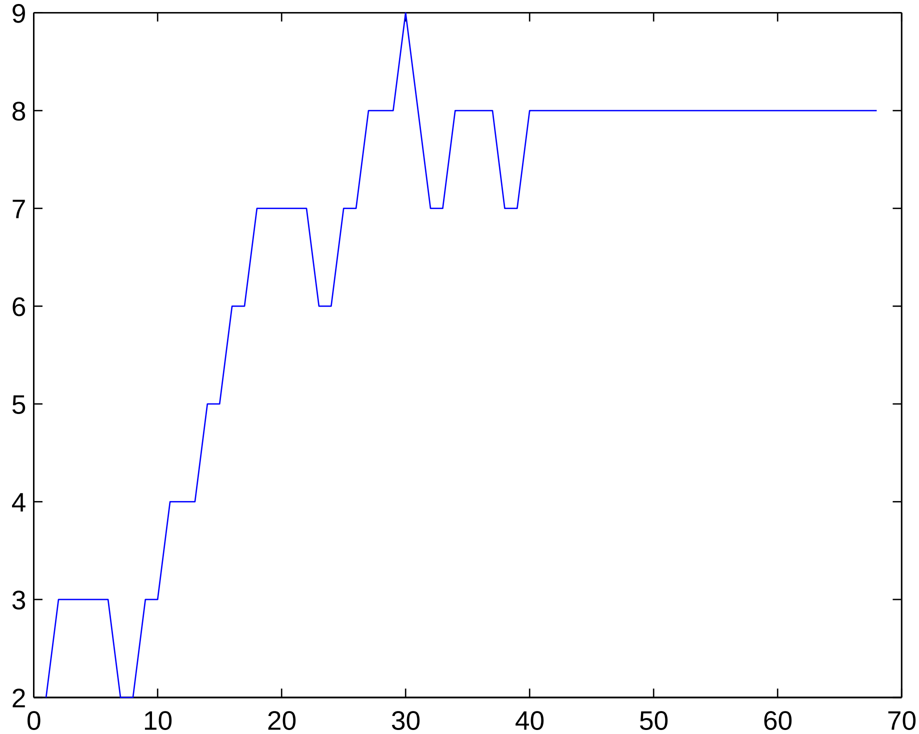

Figure 1: MAC Layer packet frame length selection with fixed MAC Layer Collision Probability = 0.5

Packet length	10	20	30	40	50	60	70	80	90	100
Throughput	0	1.15	1.23	1.53	2.16	2.59	2.56	3.00	0	0

Figure 2: Net Throughput values with fixed MAC Layer Collision Probability = 0.5. The zero values are so because the algorithm learns a negative ΔTh value for adopting these actions(from any other given action) and hence never adopts the concerned actions

References

[Bia00] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *Selected Areas in Communications, IEEE Journal on*, 18(3):535–547, 2000.