A New Theoretic Framework for Cross-Layer Optimization with Message Exchanges

Fangwen Fu and Mihaela van der Schaar
Electrical Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, USA
{fwfu, mihaela}@ee.ucla.edu

Abstract—**Existing cross-layer optimization solutions are often based on the assumption that each layer's parameters are fully accessible to other layers or a centralized optimizer, thereby violating the layered network architecture of the protocol stack. This paper presents a new systematic framework named layered Markov decision process (MDP) for cross-layer optimization, which allows each layer to make autonomous decisions individually and determine the optimal message to be exchanged among layers in order to cooperatively maximize the utility of the wireless station. Hence, this layered cross-layer framework does not change the current layered architecture and is suitable for the delay-sensitive applications over wireless networks.**

Keywords-Cross-layer optimization; layered MDP; message exchange; environmental dynamics.

I. INTRODUCTION

In layered network architectures such as the Open Systems Interconnection (OSI) model [1], the functionality of each layer is specified in terms of the services that it receives from the layer(s) below and that it is required to provide to the layer(s) above. The advantage of layered architectures is that the designer or implementer of the protocol or algorithm at a particular layer can focus on the design of that layer, without being required to consider all the parameters and algorithms of the rest of the stack. However, in current layered network implementations, each layer often optimizes its strategies and parameters individually. This generally results in sub-optimal performance for the users/applications [2]. Cross-layer optimization solutions have been proposed in recent years to improve the performance of network users operating in a time-varying, error-prone wireless environment. As pointed out in [3], these solutions optimize the protocol parameters in an integrated fashion by jointly and simultaneously considering the dynamics at each layer and requiring layers to provide access to their internal protocol parameters to other layers. However, a majority of these integrated approaches violates the layered network architecture of the protocol stack, thereby requiring a complete redesign of current networks and protocols and leading to a high implementation cost [3]. This motivates us to develop a new cross-layer optimization framework, which adheres to the current layered network architecture, and allows layers to determine their own protocol parameters, and exchange only the information they desire with other layers.

Unlike the previous works that jointly optimize the cross-layer strategies in a centralized way, we propose a layered Markov decision process (MDP) solution to drive the cross-layer optimization. In this layered MDP framework, each layer makes its transmission decision (i.e. selects the transmission strategies, e.g. packet scheduling in the application (APP) layer, retransmission in the MAC layer and modulation selection in the physical (PHY) layer) in an autonomous manner, by considering the dynamics experienced at that layer as well as the information available from other layers. Importantly, using this layered optimization framework, we do not change the current layered architecture of the protocol stack. Moreover, the current algorithms and protocols implemented at each layer also remain unaffected, as the proposed framework requires only the exchange of information across layers and the optimization of available parameters at each layer. To exchange information across multiple layers, we define a message exchange mechanism in which the content of the message captures the performed transmission strategies and experienced dynamics at each layer However, the *format* of the message is independent of the transmission strategies, protocols and dynamics implemented at each layer.

The rest of the paper is organized as follows. Section II discusses the structure of the cross-layer optimization and formulates the cross-layer design as an MDP problem. Section III presents a layered value iteration algorithm for optimally solving the layered MDP. Section IV discusses the preliminary results of the layered MDP. The paper concludes in Section V.

II. CROSS-LAYER OPTIMIZATION PROBLEM FORMULATION

A. *Structure of cross-layer optimization*

As argued in [8], many complex systems have a "nearly decomposable, hierarchical structure", with the subsystems interacting only weakly with each other. The wireless communication system is such a system and each layer can be viewed as a subsystem. This simple and elegant structure of the layered architecture is not explicitly considered in the cross-layer optimization context. Our research focuses on building a formal framework for exploring the specific layered structure of wireless networks in order to find the optimal cross-layer strategies for networked devices.

Depending on the system design [2], there can be different types of dependencies between layers. In this paper, we consider that the decision on transmission strategies at one layer only depends on its neighboring layers. However, the framework we will propose is also applicable for other dependency structures. Specifically, we consider one WSTA transmitting its time-varying traffic to another WSTA (e.g. base station) over a wireless network (e.g. wireless LAN, cellular network, etc.). We also assume that there are L participating

978-1-4244-2219-7/08/$25.00 (c)2008 IEEE

layers in the protocol stack. Each layer is indexed $l \in \{1,...,L\}$ with layer 1 corresponding to the lowest participating layer (e.g. PHY layer) and layer L corresponding to the highest participating layer (e.g. APP layer). The WSTA interacts with the dynamic environment at various layers in order to maximize the application utility.

1) State definition

In wireless communications, different states can be defined at each layer to capture the current undergoing dynamics. For example, the signal to noise ratio (SNR)[1] at the PHY layer [4], the amount of transmission opportunities acquired at MAC layer [9], and the amount of packets with various delay-deadlines at the APP layer [4] can be defined as the states of those layers. These states are assumed to be Markovian, as in [4][9]. Moreover, in this paper, the states at each layer are also defined in such a way that future transmission strategies can be determined solely based on the current state, and independently of the past history. In other words, the state encapsulates all the past information required for future strategy adaptation. Note that the state transition may be affected by the transmission strategies performed by the wireless transceiver as well as by the environmental dynamics. When considering the layered architecture of current networks, we are able to define a state $s_l \in \mathcal{S}_l$ for each layer l. Then, the state of the entire wireless user is denoted by $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$, with $\boldsymbol{\mathcal{S}} = \prod_{l=1}^{L} \mathcal{S}_l$.

2) Classsification of actions

In a layered architecture, a WSTA takes different transmission actions in each state of each layer. The transmission actions can be classified into two types at each layer l: an *external action* is performed to determine the state transition, and an *internal action* is performed to determine the QoS provided to the upper layers for the packet(s) transmission. The QoS at each layer is formally defined later in this project description. Hence, the internal actions in all the layers jointly determine how many packets can be successfully delivered to the destination. In other words, the internal actions performed at all the layers may affect the state transition in the highest layer, e.g. the APP layer. For specific applications, the layers may have only one type of action (e.g. internal action). This does not affect our proposed framework. We also note that, for some applications, the action may simultaneously determine the QoS provided to the upper layer and the state transition due to multi-user interaction in wireless networks.

The external actions at each layer l are denoted by $a_l \in \mathcal{A}_l$, where $\mathcal{A}_l$ is the set of the possible external actions available at layer l. The external actions for the WSTA in all the layers are denoted by $\boldsymbol{a} = [a_1,...,a_L] \in \boldsymbol{\mathcal{A}}$, where $\boldsymbol{\mathcal{A}} = \prod_{l=1}^{L} \mathcal{A}_l$. The internal actions are denoted by $b_l \in \mathcal{B}_l$, where $\mathcal{B}_l$ is the set of the possible internal actions available at layer l. The internal actions are performed by the WSTA to efficiently *utilize* the wireless medium given the network resource allocation and its own resource budget (e.g. power constraint), by providing the QoS required by the supported applications. The internal actions for the WSTA across all the layers are denoted by $\boldsymbol{b} = [b_1,...,b_L] \in \boldsymbol{\mathcal{B}}$, where $\boldsymbol{\mathcal{B}} = \prod_{l=1}^{L} \mathcal{B}_l$. Hence, the action at layer l is the aggregation of external and internal actions, denoted by $\xi_l = \begin{bmatrix} a_l & b_l \end{bmatrix} \in \mathcal{X}_l$, where $\mathcal{X}_l = \mathcal{A}_l \times \mathcal{B}_l$. The joint action of the WSTA is denoted by $\boldsymbol{\xi} = [\xi_1,...,\xi_L] \in \prod_{l=1}^{L} \mathcal{X}_l$.

3) Layered state transition model

The external actions drive the state transition. In this section, we examine the structure of state transition model and the underlying models for environmental dynamics. In general, because states are Markovian, the state transition of the WSTA only depends on the current state $\boldsymbol{s}$, the current performed actions, and the environmental dynamics. The corresponding transition probability is denoted by $p(\boldsymbol{s}' \mid \boldsymbol{s}, \boldsymbol{\xi})$. This global state transition can be compactly represented using dynamic decision network [10]. Formally, the transition model is presented by

$$p(\boldsymbol{s}' \mid \boldsymbol{s}, \boldsymbol{\xi}) = \prod_{l=1}^{L-1} p(s_l' \mid \text{parent}(s_l'), \text{action}(s_l')) \qquad (1)$$

where $\text{parent}(s_l')$ represents the parents of the state s_l' on which the transition of s_l' depends, and $\text{action}(s_l')$ represents the set of actions that affect the transition of s_l'. Based on the structure of actions, the transition probability can be decomposed as

$$p(\boldsymbol{s}' \mid \boldsymbol{s}, \boldsymbol{\xi}) = \prod_{l=1}^{L-1} p(s_l' \mid \boldsymbol{s}'_{1\to l-1}, s_l, a_l) p(s_L' \mid \boldsymbol{s}'_{1\to L-1}, \boldsymbol{s}, a_L, \boldsymbol{b}). \quad (2)$$

In other words, $\text{parent}(s_l') = \{s'_{1\to l-1}, s_l\}$, $action\{s_l'\} = \{a_l\}$ for $l \in \{1,...,L-1\}$ and $\text{parent}(s_L') = \{s'_{1\to L-1}, \boldsymbol{s}\}$, $action\{s_L'\} = \{a_L, \boldsymbol{b}\}$. The transition model is illustrated in Figure 1.

4) Decomposed utility function

The utility gain obtained in layer L is based on the states and internal actions at each layer and it is denoted by $g_L(\boldsymbol{s}, \boldsymbol{b})$. The transmission cost at layer l represents the cost of performing both the external and internal actions, e.g. the amount of power allocated to determine the channel conditions at the PHY layer or the cost spent to acquire the transmission opportunities at the MAC layer. In general, the transmission cost of performing the external (internal) action at layer l is denoted by $c_l(s_l, a_l)$ ($d_l(s_l, b_l)$), which is a function of the external (internal) action and the state of layer l. Based on the transition model and action structure, that the utility form is decomposed as

$$R(\boldsymbol{s}, \boldsymbol{\xi}) = g_L(\boldsymbol{s}, \boldsymbol{b}) - \sum_{l=1}^{L} \lambda_l^a c_l(s_l, a_l) - \sum_{l=1}^{L} \lambda_l^b d_l(s_l, b_l), \qquad (3)$$

where λ_l^a (λ_l^b) is a external (internal) Lagrangian multiplier in layer l, determined by the WSTA to trade off the utility gain at layer L and transmission cost at all the layers. The optimal Lagrangian multipliers depend on the available resource budget and can be obtained as in [5]. Each component of the utility function is illustrated in Figure 1.

[1] The SNR in the PHY layer can be determined based on the allocated power, the experienced channel fading and interference from other wireless transmitters. Without the knowledge of other wireless users' power allocation, the Markov assumption about the experienced SNR is reasonable and has been used in [4]. The Markov model for the MAC layer's state (i.e. the amount of time/frequency band) can be verified in the same way.

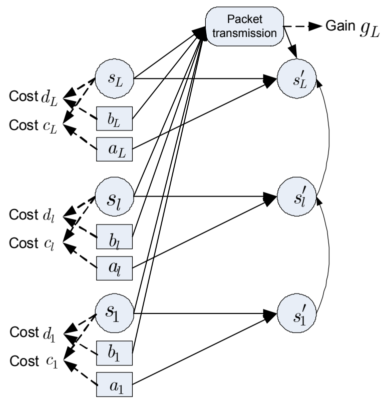


Figure 1. Layered transition model and components of utility function.

B. Problem formulation

As described above, the state transition at each layer is controlled by the external actions. For simplicity, we first assume that the state transition in each layer is synchronized and operates at the same time scale, such that the transition can be discretized into stages during which the WSTA has a fixed state and performs static actions. The length of each stage can be determined based on how fast the environment changes. We use a superscript k to denote stage k. Hence, the state of the WSTA at stage $k \in \mathbb{N}$ is denoted by $\boldsymbol{s}^k$ with each element s_l^k being the state of layer l; similarly, the joint action performed by the WSTA at state k is $\boldsymbol{\xi}^k$ with each element $\xi_l^k = \left[a_l^k, b_l^k\right]$. The state transition probability is given by Eq. (2) and the stage reward is given by Eq. (3). Specifically, we assume that the WSTA will maximize the discounted accumulative reward, which is defined as

$$\sum_{k=0}^{\infty} (\gamma)^k R(\boldsymbol{s}^k, \boldsymbol{\xi}^k \mid \boldsymbol{s}^0), \tag{4}$$

where γ is a discounted rate, with $0 \le \gamma < 1$ and $\boldsymbol{s}^0$ is the initial state.

The transmission strategies at each layer can be obtained by jointly maximizing the discounted reward defined in Eq. (4). This optimization problem can be formulated as a MDP [6].

However, the central optimizer makes decisions on the transmission action selection for all the layers, and it unfortunately violates the layered architecture since it obliges each layer to take the actions dictated by the central optimizer and the layers have no freedom to adapt their own actions to the environmental dynamics. To overcome the problems associated with the centralized cross-layer optimization, in this research we propose a layered MDP framework, which enables the layers to make optimal decisions on the transmission actions and exchange information autonomously. In this way, the layered architecture is kept unchanged.

III. Layered MDP Formulation

The problem structure discussed in Section 2 enables us to decompose the MDP into a layered MDP for the cross-layer optimization problem which is defined as follows:

Definition (**Layered MDP with information exchange**) The layered MDP model with information exchange is given by the tuple $\mathcal{M} = \left\langle \mathcal{L}, \mathcal{S}, \{\mathcal{X}_l\}_{l=1}^{L}, \{\Theta_{l,l+1}\}_{l=1}^{L-1}, \{\Theta_{l,l-1}\}_{l=2}^{L}, p, R, \gamma \right\rangle$, where

- $\mathcal{L} = \{1, ..., L\}$ is a set of L layers, each of which takes the internal and external actions individually.
- $\mathcal{S}$ is a finite set of states, each element $\boldsymbol{s} \in \mathcal{S}$ of which contains $[s_1, \cdots, s_L]$.
- $\boldsymbol{\mathcal{X}}_l$ is a finite set of actions available to layer l, each element $\xi_l \in \mathcal{X}_l$ of which contains the external and internal actions, i.e. $\xi_l = [a_l, b_l]$.
- $\Theta_{l,l+1}$ is the message set sent by layer l to its upper layer $l+1$, where $\theta_{l,l+1} \in \Theta_{l,l+1}$ represents a message sent by layer l to its upper layer $l+1$ (i.e. upward message).
- $\Theta_{l,l-1}$ is the message set sent by layer l to its lower layer $l-1$, and $\theta_{l,l-1} \in \Theta_{l,l-1}$ represents a message sent by layer l to its lower layer $l-1$ (i.e. downward message).
- p is the transition probability function. $p(\boldsymbol{s}' \mid \boldsymbol{s}, \boldsymbol{\xi})$ is the probability of moving from state $\boldsymbol{s} \in \mathcal{S}$ to the state $\boldsymbol{s}' \in \mathcal{S}$ when layer $l \in \mathcal{L}$ performs action ξ_l. We assume that the transition model is stationary and independent of the stage (i.e. time).
- $R : \mathcal{S} \times \prod_{l=1}^{L} \boldsymbol{\mathcal{X}}_l \mapsto \mathbb{R}$ is the system stage reward function which has the form of $R(\boldsymbol{s}, \boldsymbol{\xi})$, i.e. the reward is determined by the state and actions in each layer.
- γ is the discounted factor.

Upward message: At state $\boldsymbol{s}$, by deploying the internal actions, the WSTA can determine for each layer (i) the probability, $\varepsilon_l(\boldsymbol{s}_{1\to l}, \boldsymbol{b}_{1\to l})$, of a packet being dropped due to expiration; (ii) the amount of time, $t_l(\boldsymbol{s}_{1\to l}, \boldsymbol{b}_{1\to l})$, it takes to transmit a packet on average; and (iii) the cost $d_l(s_l, b_l)$ associated with its transmission. We define a QoS level as a three-tuple $Z_l = [\varepsilon_l, t_l, f_l]^T$ to represent the transmission result where $f_l = \sum_{l'=1}^{l} \lambda_{l'}^b d_{l'}(s_{l'}, b_{l'})$. The QoS at layer l represents the service layer l provides to its upper layer $l+1$. Using the QoS, layer $l+1$ does not need to know the actions and dynamics at lower layers. That is, the QoS is a sufficient statistics of the states and internal actions performed in the lower layers. By knowing QoS Z_{L-1} provided from layer $L-1$, layer L can computed the internal reward $R_{in} = g_L(\boldsymbol{s}, \boldsymbol{b}) - \sum_{l=1}^{L} \lambda_l^b d_l(s_l, b_l)$. In other words, the internal reward R_{in} is independent of the states and actions in the lower layers, given QoS Z_{L-1} provided from layer $L-1$. Hence, the upward message is $\theta_{l,l+1} = \mathcal{Z}_l$, where $\mathcal{Z}_l$ is the necessary QoS levels required by the upper layers.

Downward message: Based on our layered MDP framework, we propose a layered value iteration algorithm by allowing information exchange between adjacent layers. As defined above, each layer in the layered MDP is regarded as an autonomous entity that needs to determine its own actions. However, the layers can cooperate with each other using the information exchange in order to find the optimal transmission strategies as in the value iteration for the central MDP. By the value iteration [6], finding the optimal transmission strategies is equivalent to find the optimal state-value function $V^*(\mathbf{s})$. The value iteration is decomposed into L layered sub-value iterations.

The layered value iteration can be performed as follows: at each iteration n, layer L performs the sub-value iteration to obtain the state-value function $V^*_{n,L-1}\left(\mathbf{s}'_{1\to L\text{-}1}\right)$, which serves as future state-value function at layer $L-1$. Then, layer l performs the sub-value iteration to generate $V^*_n\left(\mathbf{s}'_{1\to l-1}\right)$ based on the future state-value function from layer $l+1$. Finally, layer 1 performs the sub-value iteration to generate the state-value function $V^*_{n,L}\left(\mathbf{s}_{1\to L}\right)$, which is $V^*_n(\mathbf{s})$, as in the centralized value iteration. Then, the downward message exchanged from layer $l+1$ to layer l is $\theta_{l+1,l} = \left\{V^*_{n-1}\left(\mathbf{s}'_{1\to l}\right)\right\}$.

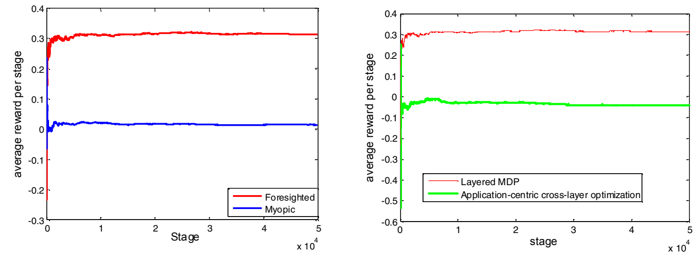
By decomposing the value iteration for the central MDP, we can obtain the following theorem.

Theorem: The state-value function $V^*(\mathbf{s})$ corresponding to the optimal policy can be obtained using a layered value iteration algorithm. At iteration n, each layer l performs a sub-value iteration to determine the optimal external and internal policies, given the downward message including $V^*_l\left(\mathbf{s}_{1\to l-1}\right)$ and optimal QoS level Z^*_l from layer $l+1$.

The proof can be found in [7] and is omitted due to the space limitation.

IV. Preliminary Results

We now show the preliminary results that highlight the performance improvement of our layered MDP framework compared to the myopic cross-layer optimization and the cross-layer optimization with ad-hoc message exchanges. We consider the optimization of the transmission strategies available at the APP, MAC, and PHY layers, i.e. $L=3$. In the PHY layer, the SNR can be modeled as a finite state Markov chain (FSMC) [4]. To satisfy the service requirement from upper layers, the PHY layer adapts its modulation and channel coding schemes based on the current SNR. In the MAC layer, the wireless user acquires the transmission opportunities from the shared network. Besides the resource acquisition, the MAC can also perform error control algorithms (e.g. ARQ) to improve the service provided to the upper layers. In the APP layer, the WSTA generates delay-sensitive video data. The number of packets available for transmission depends on the source coding parameters adaptation as well as the transmission strategies at the lower layers. We first compare the performance of the myopic cross-layer optimization (i.e. $\gamma = 0$) versus our proposed foresighted cross-layer optimization. Figure 2 (a) shows the average reward per stage for both the myopic policy and foresighted policy. The average reward obtained by the foresighted policy is 0.3115, while the average reward by the myopic policy is only 0.0132. The simulation results demonstrate that the foresighted policy can achieve much better performance (approximately 24 times better in this simulation) than the myopic policy. To compare the performance loss for traditional cross-layer optimization with the ad-hoc message exchange, we also depict the average reward per stage for the layered MDP and application-centric cross-layer optimization in Figure 2 (b). From Figure 2 (b), we note that the ad-hoc message exchange result in suboptimal performance.



(a) (b)

Figure 2. (a) Average reward per state for myopic cross-layer optimization and foresighted cross-layer optimization; (b) Average reward per state for formal message exchange in layered MDP and application-centric cross-layer solution.

V. Conclusions

In the proposed algorithm, the optimization of solving the optimal actions are decomposed into layered optimization sub-problems each of which corresponds to a sub-value iteration. First, each layer is not required to know the dynamics model and possible internal and external actions from other layers, but only its own dynamics and actions. Second, the format of the messages between layers is independent of the protocols deployed in each layers although the content of the messages characterizes the dynamics and performed actions at each layer. The layered framework allows us to analyze how the limited message exchanges among layers affect the WSTA's utility and what is the bound by considering all possible environmental dynamics.

References

[1] D. Bertsekas *et al*, "Data networks," Prentice Hall, Inc. Upper Saddle River, NJ, 1987.

[2] M. van der Schaar *et al*, "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms," *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, Aug. 2005.

[3] V. Kawadia *et al*, "A cautionary perspective on cross-layer design," *IEEE Wireless Commun.*, pp. 3-11, vol. 12, no. 1, Feb. 2005.

[4] T. Holliday *et al*, "Optimal Power Control and Source-Channel Coding for Delay Constrained Traffic over Wireless Channels," *Proceedings of ICC'02*, May 2002.

[5] D. Djonin *et al*, "MIMO transmission control in fading channels-a constrainted Markov decision process formulation with monotone randomized policies," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5069-5083, Oct. 2007.

[6] D. P. Bertsekas, "Dynamic programming and optimal control," 3rd, Athena Scientific, Belmont, Massachusetts, 2005.

[7] F. Fu *et al*, "A new theoretic foundation for cross-layer optimization," Technique report, Dec. 2007, available on http://xxx.arxiv.org/abs/0712.2497.

[8] H. Simon, "The architecure of complexity," *Proceedings of the American Philosophical Society*, vol. 106, no. 6, Dec. 1962.

[9] S. Shankar, C.T. Chou, K. Challapali, and S. Mangold, "Spectrum agile radio: capacity and QoS implications of dynamic spectrum assignment," *Global Telecommunications Conference,* Nov. 2005.

[10] T. Dean, K. Kanazawa, "A model for reasoning about persistence and causation," *Computational Intelligence*, 5(3), 142–150, 1989.

978-1-4244-2219-7/08/$25.00 (c)2008 IEEE