A NEW THEORETIC FRAMEWORK FOR CROSS-LAYER OPTIMIZATION

Fangwen Fu, and Mihaela van der Schaar

Multimedia Communications and Systems Lab, Electrical Engineering Department,
University of California Los Angeles (UCLA)
{fwfu, mihaela}@ee.ucla.edu

ABSTRACT

Cross-layer optimization aims at improving the performance of network users operating in a time-varying, error-prone wireless environment. However, current solutions often rely on ad-hoc optimization approaches, which ignore the different environmental dynamics experienced at various layers by a user and violate the layered network architecture of the protocol stack. This paper presents a new theoretic framework in which the cross-layer optimization problem is formulated as a layered Markov decision process (MDP). In this framework, each layer adapts its own protocol parameters and exchanges information (messages) with other layers in order to cooperatively maximize the performance of the wireless user. Hence, this layered cross-layer framework does not change the current layered architecture and is suitable for the delay-sensitive applications over wireless networks.

Index Terms— Cross-layer optimization, layered MDP, information exchange, environmental dynamics.

1. INTRODUCTION

To optimize the different protocol parameters from different layers in the Open Systems Interconnection (OSI) model [1], the wireless stations (WSTAs) need to consider the dynamic wireless network "environment" resulting from the repeated interaction with other stations, the experienced time-varying channel conditions and, for delay-sensitive applications, the time-varying source characteristics. Moreover, it should be noted that a WSTA needs to jointly optimize the selected protocol parameters within each layer such that the utility of the WSTA is maximized. The joint optimization of the transmission strategies at the various layers is referred to as cross-layer design [2][3].

The advantage of the current layered architecture is that the designer or implementer of the protocol or algorithm at a particular layer can focus on that layer without worrying about the rest of stack [3]. However, most existing cross-layer design solutions advocate improving the system utility by violating the current layered architecture of wireless networks. These cross-layer interactions create the dependencies among the layers which will affect not only the concerned layer but also other layers. Hence, such

solutions are undesirable because they require a complete redesign of current networks and protocols and thus, require a high implementation cost [3].

Furthermore, some existing cross-layer design solutions aim at maximizing the WSTA's utility by jointly adapting the transmission strategies across multiple layers to the current environmental dynamics [2]. These solutions, however, neglect that the environmental dynamics are also affected by the cross-layer transmission strategies, thereby affecting the future utility derived by the network users.

Unlike the previous works that jointly optimize the cross-layer strategies in a centralized way, we propose a layered 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 currently 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 2 discusses the problem settings for the cross-layer optimization and formulates the cross-layer design as an MDP problem. Section 3 presents a layered value iteration algorithm for optimally solving the layered MDP. Section 4 discusses the advantage of the layered MDP. The paper concludes in Section 5.

2. CROSS-LAYER PROBLEM STATEMENT

We consider one WSTA transmitting its time-varying traffic (e.g. multimedia) 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 layers[1] 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.

2.1 States

In this paper, the state of the layers is defined such that future transmission strategies can be determined independent of the past history given the current state. In other words, the state encapsulates all the past information required for future strategy adaptation. We refer to this type of state as Markovian state. 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 WSTA is denoted by $s \in \mathcal{S}$, with $\mathcal{S} = \prod_{l=1}^{L} \mathcal{S}_l$.

2.2 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 service (i.e. QoS) provided to the upper layers for the packet(s) transmission.

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 $a = [a_1,...,a_L] \in \mathcal{A}$, where $\mathcal{A} = \prod_{l=1}^{L} \mathcal{A}_l$. 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 for the WSTA across all the layers are denoted by $b = [b_1,...,b_L] \in \mathcal{B}$, where $\mathcal{B} = \prod_{l=1}^{L} \mathcal{B}_l$. 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 $\xi = [\xi_1,...,\xi_L] \in \prod_{l=1}^{L} \mathcal{X}_l$.

Due to the layered architecture of the wireless network, the state transition probability in this paper can be decomposed as follows.

$$p(s' \mid s, \xi) = \prod_{l=1}^{L-1} p(s_l' \mid s_{1\rightarrow l-1}', s_l, a_l) p(s_L' \mid s_{1\rightarrow L-1}', s, a_L, b) \quad (1)$$

where $s_{1\rightarrow l}' = [s_1',...,s_l']$.

This decomposition is due to the layered network architecture and enables us to develop a layered MDP framework, which will be presented in Section 3.

[1] If one layer does not participate in the cross-layer design, it can simply be omitted. Hence, we consider here only the L participating layers.

2.3 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 $g(s, b)$. 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 or the tax (tokens, money) spent for consuming wireless resources. 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. For illustration, we assume that the reward is defined as

$$R(s, \xi) = g(s, b) - \sum_{l=1}^{L} \lambda_l^a c_l(s_l, a_l) - \sum_{l=1}^{L} \lambda_l^b d_l(s_l, b_l) \quad (2)$$

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 and transmission cost. We assume that the Lagrangian multipliers λ_l^a and λ_l^b are known. The optimal Lagrangian multipliers depend on the available resource budget and can be obtained as in [6].

2.4 Foresighted decision making

As described in Section 2.2, the state transition at each layer is controlled by the external actions. For simplicity, we 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 constant state and performs static actions. We use a superscript k to denote stage k.

Unlike the tradition cross-layer adaptation that focuses on the myopic (i.e. immediate) utility, in the proposed cross-layer framework, the goal is to find the optimal internal and external actions at each stage such that a cumulative function of the rewards is maximized. We refer to this decision process as the *foresighted* cross-layer decision. By maximizing the cumulative reward, the WSTA is able to take into account the impact of the current actions on the future reward.

Specifically, we assume that the WSTA will maximize the discounted accumulative reward, which is defined as

$$\sum_{k=0}^{\infty} (\gamma)^k R(s^k, \xi^k \mid s^0) \quad (3)$$

where γ is a discounted rate with $0 \leq \gamma < 1$ and s^0 is the initial state.

3. LAYERED MDP FORMULATION

Existing cross-layer optimization frameworks require a central controller to decide the parameter configuration assuming that the complete information from all the layers is available to the central controller [7]. The foresighted cross-layer optimization can be formulated as an MDP and solved using value iteration [8]. To solve this central MDP problem, the central optimizer needs to know the following:
• the state space at each layer;

- the action space at each layer;
- probability distribution describing the state transition (i.e. environmental dynamics);
- state reward function of the states and performed actions;

Although the centralized cross-layer optimization (formulated as the MDP problem) considers the information exchange among the layers in an indirect way, it unfortunately violates the layered architecture. However, 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 $s \in \mathcal{S}$ of which contains $[s_1, \cdots, s_L]$.
- $\mathcal{X}_l$ 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. $p(s' \mid s, \xi)$ is the probability of moving from state $s \in \mathcal{S}$ to the state $s' \in \mathcal{S}$ 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).
- $R: \mathcal{S} \times \prod_{l=1}^L \mathcal{X}_l \mapsto \mathbb{R}$ is the system stage reward function which has the form of $R(s, \xi)$, i.e. the reward is determined by the state and actions in each layer.
- γ is the discounted factor.

The framework of the layered MDP with information exchange for the foresighted cross-layer optimization problem is illustrated in Figure 1. From this figure, we observe that the layer optimizer is not required to know other layers' state space, action space and dynamics models.

Upward message: At state s^k, by deploying the internal actions, the WSTA can determine for each layer (i) the probability of the packet being successfully received at the destination; (ii) the amount of time it takes to transmit on average; and (iii) the cost associated with its transmission. The transmission result of whether a packet is successfully received, is represented by the average packet loss ratio (PLR) at layer l at stage k, which is denoted by

$\varepsilon_l^k \left(s_{1 \to l}^k, b_{1 \to l}^k \right)$ where $s_{l' \to l''}^k = \left[s_{l'}^k, ..., s_{l''}^k \right]$ and $b_{l' \to l''}^k = \left[b_{l'}^k, ..., b_{l''}^k \right]$ with $l' \leq l''$. The average amount of time spent on transmitting one packet at layer l at stage k is denoted by $t_l^k \left(s_{1 \to l}^k, b_{1 \to l}^k \right)$. The aggregated transmission cost incurred by performing internal actions at layer l is defined by $f_l^k \left(s_{1 \to l}^k, b_{1 \to l}^k \right) = \sum_{l'=1}^l \lambda_{l'}^b d_{l'} \left(s_{l'}^k, b_{l'}^k \right)$.

To compute the reward function associated with the internal actions, i.e. $R_{in}^k = g \left(s^k, b^k \right) - \sum_{l=1}^L \lambda_l^b d_l^k \left(s_l^k, b_l^k \right)$, layer L has to know the packet loss probability, the average amount of time for packet transmission and the internal transmission cost provided from the lower layers in stage k. We can define a message which captures this information from lower layers. This message is the QoS at layer l which is defined as a three-tuple $Z_l^k = \left[\varepsilon_l^k, t_l^k, f_l^k \right]^T$. 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.

By knowing QoS Z_{L-1}^k provided from layer $L-1$, layer L can computed R_{in}^k. In other words, the internal reward R_{in}^k is independent of the states and actions in the lower layers, given QoS Z_{L-1}^k provided from layer $L-1$. Hence, the upward message is $\theta_{l,l+1} = \mathcal{Z}_l^k$ where $\mathcal{Z}_l^k$ is the necessary QoS levels required by the upper layers. The more details about the upward message are presented in [9].
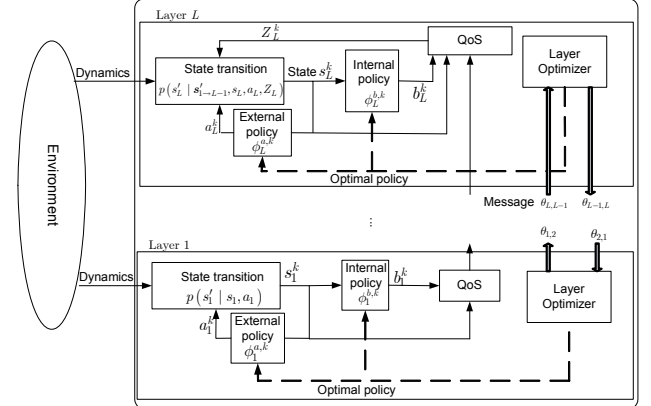

Figure 1. Layered cross-layer optimization framework

Downward message: As in the definition of the layered MDP, each layer is regarded as an autonomous entity that performs its own actions. However, the layers can cooperate via the information exchange to find the optimal state-value function $V^*(s)$ as in the value iteration for the central MDP [8]. By decomposing the value iteration for the central MDP given in [9], we can obtain the following theorem.

Theorem 1: The state-value function $V^*(s)$ corresponding to the optimal policy can be obtained using a layered value iteration algorithm. At iteration n, each layer performs a sub-value iteration which is given in Table 1.

The proof is omitted here due to the space limitation and can be found in [9].

Table 1. Sub-value iteration at each layer.

Layer	Sub-value iteration form at iteration n	
L	$V_{n,L-1}^{*}\left(s_{1\to L-1}'\right) = \max_{a_L\in\mathcal{A}_L, Z_L\in\mathcal{Z}_L}$ $\begin{bmatrix} R_{in}\left(s_L,Z_L\right) - \lambda_L c_L\left(s_L,a_L\right) + \\ \gamma\sum_{s_L\in\mathcal{S}_L} p\left(s_L'\mid s_{1\to L-1}',s_L,a_L,Z_L\right)V_{n-1,L}^{*}\left(s_{1\to L}'\right) \end{bmatrix}$	(4)
l	$V_{n,l-1}^{*}\left(s_{1\to l-1}'\right) = \max_{a_l\in\mathcal{A}}$ $\left[-\lambda_l c_l\left(s_l,a_l\right) + \sum_{s_l'\in\mathcal{S}_l} p\left(s_l'\mid s_{1\to l-1}',s_l,a_l\right)V_{n,l}^{*}\left(s_{1\to l}'\right) \right]$	(5)
1	$V_{n,L}^{*}\left(s_{1\to L}\right) = \max_{a_1\in\mathcal{A}_1}$ $\left[-\lambda_1 c_1\left(s_1,a_1\right) + \sum_{s_1'\in\mathcal{S}_1} p\left(s_1'\mid s_1,a_1\right)V_{n,1}^{*}\left(s_1'\right) \right]$	(6)

The layered value iteration is performed as follows: at each iteration n, layer L performs the sub-value iteration as in Eq. (4) to obtain the state-value function $V_{n,L-1}^{*}\left(s_{1\to L-1}'\right)$ which services as future state-value function at layer $L-1$. Then, in general, layer l performs the sub-value iteration as in Eq. (5) based on the future state-value function from layer $l+1$ to generate $V_n^{*}\left(s_{1\to l-1}'\right)$. Finally, layer 1 performs the sub-value iteration as in Eq. (6) to generate the state-value function $V_{n,L}^{*}\left(s_{1\to L}\right)$, which is $V_n^{*}(s)$ as in the centralized value iteration.

Then the message exchanged from layer $l+1$ to layer l is $\theta_{l+1,l} = \left\{V_{n-1}^{*}\left(s_{1\to l}'\right)\right\}$. The upward and downward message exchanges are presented in Table 2.

Table 2. Message exchanges between layers at iteration n.

Layer	Upward Message $\theta_{l,l+1}$	Downward Message $\theta_{l,l-1}$
L	$\varnothing$	$\left\{V_n^{*}\left(s_{1\to L-1}\right)\right\}$
l	$\mathcal{Z}_l$	$\left\{V_n^{*}\left(s_{1\to l-1}\right)\right\}$
1	$\mathcal{Z}_1$	$\varnothing$

4. ADVANTAGES OF LAYERED VALUE ITERATION

In this section, we highlight the advantages of the proposed layered value iteration for the layered MDP, compared to the centralized value iteration as in [9].

In the central MDP, the central optimizer is required to completely know dynamics model (i.e. states, transition probability) and possible internal and external actions of all the layers which are protocol-dependent. Hence, the mechanism of information exchange between the central optimizer and the layers is also protocol-dependent. However, in the proposed algorithm, the optimization of solving the optimal actions are decomposed into layered optimization sub-problems each of which corresponds to a value iteration shown in Table 1. 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 listed in Table 2 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. This layered framework is essential for the design of time-varying delay-sensitive applications over wireless networks.

5. CONCLUSION

In this paper we formulated the dynamic cross-layer optimization problem as a layered MDP with information exchanges among layers. For the layered MDP, we design a layered value iteration algorithm and message exchange mechanism among layers. The layered cross-layer framework does not break up the current layered network architecture and hence, is suitable for the time-varying, delay-sensitive applications (e.g. multimedia) over dynamic wireless networks.

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] A. Ekbal *et al*, "QoS-constrained physical layer optimization for correlated flat-fading wireless channels," *proceedings of ICC'04*, vol. 7, pp. 4211-4215, June., 2004.

[6] 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.

[7] X. Wang *et al*, "Analyzing and optimizing adaptive modulation coding jointly with ARQ for QoS-guaranteed traffic," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, Mar. 2007.

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

[9] 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.

[10] R. S. Sutton *et al*, "Reinforcement learning: an introduction," Cambridge, MA:MIT press, 1998.

[11] Q. Zhang *et al*, "Finite-state Markov Model for Reyleigh fading channels," *IEEE Trans. Commun.* vol. 47, no. 11, Nov. 1999.

[12] F. Fu *et al*, "Non-collaborative resource management for wireless multimedia applications using mechanism design," *IEEE Transaction on Multimedia*, vol. 9, no. 4, pp. 851-868, Jun. 2007.