Cross-Layer Design in Wireless Mesh Networks

Ian F. Akyildiz, *Fellow, IEEE*, and Xudong Wang, *Member, IEEE*

***Abstract*—The conventional layered-protocol architecture does not provide optimal performance for wireless mesh networks (WMNs). The method of optimization decomposition of the protocol stack can achieve optimal network performance. This method usually results in a clean-slate protocol architecture that is different from the protocol architecture of WMNs. Such a difference actually demonstrates the need for a cross-layer design. Specific features pertaining to WMNs also show the need for cross-layer optimization across different protocol layers. In this paper, motivations for cross-layer design in WMNs are stated first. Moreover, cross-layer optimization schemes and algorithms between different protocol layers are investigated with an objective of shedding light on open research problems and new approaches. Guidelines for carrying out cross-layer design in WMNs are also provided in this paper.**

***Index Terms*—Control, cross-layer design, medium access, optimization decomposition, routing, wireless mesh networks.**

I. Introduction

A WIRELESS mesh network (WMN) consists of mesh routers and mesh clients forming a multihop wireless network [1]. It is usually connected to the Internet to provide users with backhaul access. In many application scenarios, WMNs integrate both *ad hoc* and infrastructure operation modes and interwork with other wireless networks. A WMN has many features that are much different from a wireless sensor or a mobile *ad hoc* network [1]. Furthermore, it is more concerned with scalable end-to-end throughput and satisfactory quality of service (QoS) to deliver heterogeneous traffic. Thus, it is more critical to optimize the overall network performance of WMNs across multiple protocol layers.

Whether layered-protocol design or cross-layer design is a better option to optimize protocol performance in WMNs is still an on-going research topic. The methodology of layered-protocol design carries several advantages from a protocol-transparency perspective. For example, protocols in one layer can be designed, enhanced, or even replaced without any impact on other protocol layers. However, such a methodology does not provide a mechanism for performance optimization between different protocol layers, which can significantly compromise network performance. This is particularly true for WMNs because it demands scalable network performance but is exposed to many challenging problems such as heterogeneous QoS constraints, multihop wireless communications, and variable link capacity.

Researchers have proposed to consider new protocol layering by decomposing the overall network-performance optimization [2]. As long as the optimization decomposition is successfully done, each protocol in each layer works as an optimal module to achieve the best network performance; as a consequence, cross-layer optimization has been considered in protocol layering. However, to carry out optimization decomposition, there remain many unresolved issues. A typical example is the lack of a model that can capture the stochastic dynamics in different time scales such as packet, session, connection, and topology levels.

On the other hand, the protocol layering by optimization decomposition does not necessarily match the existing protocol stack that is widely adopted in WMNs. Depending on different representation of the optimization problem, different decomposition structures may be derived, which may result in different architectures for protocol layering. However, the protocol stack in WMNs follows the classical TCP/IP protocol stack. The mismatch between the classical TCP/IP protocol stack protocol and the new protocol layering based on optimization decomposition actually illustrate that cross-layer design is highly desired if the optimum network performance must be achieved.

Therefore, it is reasonable to believe that cross-layer optimization will continue to be one of the most important tasks in protocol design for WMNs. However, critical issues must be considered for cross-layer design [3], because it has risks due to loss of protocol-layer abstraction, incompatibility with existing protocols, unforeseen impact on the future design of the network, and difficulty in maintenance and management. Thus, certain guidelines need to be followed when carrying our cross-layer design.

The rest of this paper is organized as follows. The necessity of cross-layer design is analyzed in Section II, where the challenging issues that are particular to WMNs are also discussed. In Section III, we study cross-layer design schemes and algorithms between different protocol layers. This paper is concluded in Section IV, where guidelines are provided for cross-layer design in WMNs.

II. Motivations for Cross-Layer Design

Cross-layer design has been widely used to improve the network performance, particularly in a wireless network [4]. In this section, we illustrate the need for cross-layer design in WMNs from two aspects: theoretical framework and practical factors in protocol design.

Manuscript received June 5, 2007; revised September 24, 2007 and October 16, 2007. The review of this paper was coordinated by Dr. W. Zhuang.

I. F. Akyildiz is with the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ian@ece.gatech.edu).

X. Wang is with Kiyon, Inc., San Diego, CA 92121 USA (e-mail: wxudong@ieee.org).

Digital Object Identifier 10.1109/TVT.2007.911615

A. Theoretical Framework on Layered Versus Cross-Layer Design

1) Layering as Optimization Decomposition: Layered-protocol architecture is one of the most important factors that has made networking so successful. However, there has been a lack of a systematic approach to analyze whether layering of protocols is optimal or not. The "layering as optimization decomposition" fills a gap between theoretical methods and practical aspects of protocol design. In this method, various protocol layers are integrated into one single coherent theory, in which asynchronous distributed computation over the network is applied to solve a global optimization problem in the form of generalized network utility maximization (NUM). The key idea of "layering as optimization decomposition" is to decompose the optimization problem into subproblems, each corresponding to a protocol layer and functions of primal or Lagrange dual variables, coordinating these subproblems correspond to the interfaces between layers [2]. Since different decompositions result in different layering schemes, conditions under which layering incurs no loss of optimality need to be studied, as well as the sensitivity of a layering scheme. These conditions and sensitivity can help to identify the performance differences between different layering schemes.

The basic NUM is usually formulated for protocol-layer-performance optimization, while generalized NUM needs to capture the entire protocol stack. A possible formulation of a generalized NUM is given in [2]

$$\begin{aligned} \text{maximize} \quad & \sum_s U_s(x_s, P_{e,s}) + \sum_j V_j(w_j) \\ \text{subject to} \quad & \mathbf{R}\mathbf{x} \leq \mathbf{c}(\mathbf{w}, \mathbf{P}_e) \\ & \mathbf{x} \in \mathcal{C}_1(\mathbf{P}_e),\ \mathbf{x} \in \mathcal{C}_2(\mathbf{F}),\ \text{or } x \in \Pi \\ & \mathbf{R} \in \mathcal{R},\ \mathbf{F} \in \mathcal{F},\ \mathbf{w} \in \mathcal{W}. \end{aligned} \tag{1}$$

This NUM tries to maximize the user-utility function $U(\cdot)$ and resources $V_j(\cdot)$ on the network element j. x_s and w_j denote the rate for source s and the physical layer resources at network element j, respectively. $\mathbf{R}$ is a routing matrix, and $\mathbf{x}$ denotes the link capacity as a function of physical-layer resource $\mathbf{w}$ and the desired error probability $\mathbf{P}_e$ after decoding. All physical-layer factors, such as interference, power control, etc., should be captured in function $\mathbf{c}$. Thus, the first constraint in the earlier NUM represents the behavior perceived at the routing layer. The coding and error-control mechanisms versus the rate are captured in function $\mathcal{C}_1(\cdot)$, while the contention-based MAC or scheduling-based MAC is captured in $\mathbf{C}_2(\cdot)$ and Π, respectively, where $\mathbf{F}$ is the contention matrix and Π is a schedulability constraint set. Thus, the second line of constraints stands for link-layer behavior that has taken into account the effect of the physical layer. From the earlier generalized NUM, we can see that the network performance is to be optimized at the transport layer subject to the constraints in routing, MAC, and physical layers.

The above formulation is based on a deterministic fluid model, which cannot capture the packet-level details and microscopic queuing dynamics. Thus, stochastic NUM is a preferred formulation [2]. Stochastic NUM has been an active research area, in which many challenging issues still remain to be resolved.

Whether it is a deterministic or stochastic generalized NUM, the optimization decomposition is usually carried out by the following three steps.

1) The generalized NUM is formulated independent of layering.
2) A modularized and distributed solution is developed to perform optimization by following a particular decomposition.
3) The space of different decompositions is explored such that a choice of a layered-protocol stack is made.

In the generalized NUM, the objective function is usually comprised of two parts: user- and operator-objective functions. These two parts can be integrated via a weighted sum. Another option is the multiobjective optimization that characterizes the Pareto-optimal tradeoff between user and operator objectives. Game theory can also be used to formulate the NUM with both user- and operator-objective functions.

The optimization decomposition for the generalized NUM is comprised of both horizontal and vertical decomposition.

1) Vertical decomposition: Here, the entire network functionalities are decoupled into different modules such as congestion control, routing, scheduling, MAC, power control, error control, and so on. Different modules can be classified into different layers in the protocol stack.
2) Horizontal decomposition: This aims at devising a distributed computation solution to individual module. More specifically, this step will work out a specific distributed mechanism and algorithm for protocols such as congestion control, scheduling, MAC, and so on.

As shown in [2], the optimization decomposition lays a theoretical ground for cross-layer design.

1) Optimization decomposition gives a better insight to existing layered protocols. For example, comparing a decomposition result with the existing protocol stack can tell us which layers need cross-layer optimizations and how to optimize the interactions between layers.
2) Optimization decomposition provides a systematic approach for the design of an optimized protocol architecture. Under this architecture, optimization between layers has already been considered [2], and thus, minimum efforts for cross-layer design are needed. However, such a clean-slate protocol architecture usually does not match an existing protocol stack, e.g., the widely accepted TCP/IP protocol stack for WMNs. The architecture mismatch between the optimal decomposition and an existing protocol stack indicates that cross-layer design is necessary for networks based on conventional protocol layering.
3) Optimization decomposition does not eliminate the need for cross-layer design. For example, vertical decomposition separates functionalities into different modules in different layers. However, the decomposition may still keep coupling between layers or modules. Such coupling actually proves the natural need for cross-layer design in a network.

2) Multihopping is Order-Optimal: Independent from the work of "layering as optimization decomposition," the scaling laws of transport capacity of wireless multihop networks studied in [5] also suggest that layered design is optimal.

Given a planar network in which two nodes are separated with a distance ρ_{ij}, if node i transmits a signal level of $X_i(t)$, then its received-signal level is

$$Y_i(t) = \sum_{j \neq i} \frac{e^{-\gamma \rho_{ij} X_j(t)}}{\rho_{ij}^{\delta}} + Z_i(t)$$

where $Z_i(t)$ is Gaussian noise, constant δ is the path-loss exponent, and γ is the absorption constant. In [5], the following results have been achieved.

1) The scenario of exponential attenuation: Suppose that absorption exists in the medium (i.e., $\gamma > 0$) or the path-loss exponent ρ_{ij} is larger than three, then the transport capacity, defined as the distance weighted sum of rates, grows as $\Theta(n)$, i.e., the transport capacity grows on the order of n. Furthermore, if the traffic load on each node can be balanced, then the multihop forward-and-decode strategy, treating interference as noise, is order-optimal with respect to the transport capacity.
2) The scenario of low attenuation: If $\gamma = 0$ or the path-loss exponent is small (e.g., $\delta < 3/2$), then the attenuation is low. In this scenario, other strategies like coherent multistage relaying with interference subtraction can be order-optimal with respect to the scaling law of transport capacity. This result suggests that a new protocol architecture rather than a conventional layered structure is probably needed for information transport.

In WMNs, the normal scenario is actually the exponential attenuation. Based on the result of order-optimal multihopping, it is stated in [3] that the decode-and-forward strategy can achieve optimal performance, within a constant, with regard to the network capacity. It is also pointed out in [3] that a natural way of implementing the decode-and-forward strategy is the layered-protocol architecture. Consequently, it is concluded that the cross-layer design can only improve throughput by at most a constant factor and that an unbounded performance improvement cannot be achieved.

However, such a statement can be too strong in many scenarios, particularly when we are interested in actual protocol design rather than carrying out an asymptotic analysis. As explained as follows, the results in [3] and [5] do not really prove that the cross-layer design is not necessary.

1) The theoretical results are only based on simplistic network models and only meaningful asymptotically. For a realistic wireless network, due to reasonable network size (not approaching infinity) and nonideal network models, the asymptotic scaling law does not really reflect the actual network-capacity bound. Considering cross-layer design versus layered design, their actual network capacity can be significantly different, even though the asymptotic capacity remains the same.
2) The decode-and-forward strategy does not actually imply a layered-protocol design. Almost all existing multihop wireless networks are designed based on decode-and-forward strategy, but we still see many examples of cross-layer design for improving network performance. For example, the existing protocol stack adopted in 802.11 WMNs is definitely based on a decode-and-forward strategy, but carrying out MAC/physical or MAC/routing cross-layer design is a common technology to improve network performance.

B. Features Demanding Cross-Layer Design

Several characteristics pertaining to WMNs make cross-layer design more indispensable for WMNs than that in other multihop wireless networks such as mobile *ad hoc* or wireless-sensor networks.

1) No clean-slate protocol architecture: By optimization decomposition, a new protocol architecture that is quite different from the existing standard protocol stack can also result. The well-known TCP/IP protocol stack has been widely adopted for most applications of WMNs. Thus, how to make the layered-protocol architecture derived from optimization decomposition and the TCP/IP protocol stack match with each other is a technical challenge. It is highly possible that no match can be achieved in several cases. Thus, in order to further improve the network performance without abandoning the TCP/IP protocol stack, the cross-layer design becomes indispensable.
2) Advanced physical-layer technologies: Many advanced physical-layer technologies have been adopted for WMNs in order to support applications that have high bandwidth demand. These technologies fall into several major categories.
   a) Multirate-transmission technology: This is achieved by having multiple options of modulation, coding, and power-control schemes. Different transmission rate usually results in different transmission range and interference range. With multirate-transmission technology, the same physical layer can support a different transmission rate, depending on the link quality and the environment. In a single-hop wireless network, link-adaptation protocols, which are a type of simple cross-layer design schemes, can satisfy the need for maximizing throughput. In WMNs, however, merely the link adaptation is not enough, since links within multiple hops are related to each other. Thus, in WMNs, link adaptation becomes networkwide rather than a one-hop mechanism. Thus, link adaptation is inevitably cross-related to routing and topology control. Such cross-relationship between different protocols reflects the necessity of cross-layer design.
   b) Advanced antenna technology: Directional antennas and the advance versions, such as smart antennas, can significantly reduce interference between nodes that are close to each other. Such techniques certainly

increase the network capacity but also require additional algorithms in upper layers to coordinate the antenna direction or beamforming. In a single-hop wireless network, a control algorithm located in the MAC layer, i.e., MAC/physical cross-interaction, is enough. However, in WMNs, routing needs to be considered together, since different beamforming or antenna direction impacts the routing path and *vice versa*. In other words, routing, MAC, and physical layers all need to work together. A more advanced antenna technology is multiple input and multiple output (MIMO). In a node using MIMO, advanced signaling processing technology is employed to achieve an optimal balance between link reliability and link capacity. MIMO on a point-to-point or point-to-multipoint setup has been well researched. However, taking advantage of MIMO in WMNs usually requires a networkwide-scheduling scheme.

c) Multichannel or multiradio technology: Multichannel operation (either single- or multiple-radio) can significantly reduce the interference between nodes in a multihop network. To utilize such a technology, an additional algorithm (dynamic channel allocation) must be developed in the MAC layer. This algorithm also needs to be aware of the interference from external networks. Since varying channels in different hops potentially impact the optimal routing path that can be selected, both MAC and routing protocols must work together to take advantage of the multichannel technology.

It should be noted that the above three classes of physical-layer technologies are usually integrated, which further intensify the challenge in protocol design in upper layers. For example, the multirate transmission can happen in a physical layer using MIMO and multichannel operation. For a WMN with so many advanced physical-layer features, it is more challenging to reoptimize both MAC and routing protocols.

3) Imperfect MAC: MAC has always been a critical part in all wireless networks. Many solutions are available. However, none of them is perfect because of the following two major factors: 1) The wireless medium is always imperfect in nature, and 2) the MAC itself has no guaranteed performance. In the second factor, a typical example is CSMA/CA, which is a best effort protocol and cannot provide any guarantee for delay, collisions, etc. Such unpredictable performance of the MAC can severely limit the performance of a routing protocol. For example, routing messages may not be able to send out in a congested CSMA/CA-based WMN, which in turn impacts the capability of a routing protocol. This issue is even worse in WMNs, because the performance of MAC is not just a matter of single-hop networking but multihop. Research can be carried out to constantly improve the MAC protocols for WMNs. However, as a matter of fact, if routing is not taken into account, optimal performance can only be achieved locally. Consequently, in order to achieve the ultimate goal of perfect MAC, routing must be considered as an integral part of MAC. In this sense, MAC and routing protocols in WMNs are so closely related that they should be put together as two modules in one layer or even just one module in the same protocol layer. A typical example is the upcoming IEEE 802.11s standard for 802.11 WMNs, in which MAC and routing have been put together into the same MAC layer. However, we have also noticed that the optimal interactions between MAC and routing have not been exploited yet in IEEE 802.11s.

4) Mixed traffic types with heterogeneous QoS: WMNs are expected to support a large variety of services that consist of many traffic types with heterogeneous QoS requirements. In order to deliver such services in WMNs, transport layer, routing, and MAC protocols need to cooperate smoothly; otherwise, either service quality is not ensured or the network resources may be wasted. For example, it is always preferable to use separate transport-layer protocols for VoIP, video, and data traffic. For VoIP and video traffic, finding a reliable routing path is obviously not the goal, since a path does not guarantee the quality of VoIP or video, no matter how reliable the path can be. Thus, finding a routing path must consider bandwidth allocation. This problem has been researched as a QoS-routing topic. However, when more advanced physical-layer technologies are considered, it becomes more than a QoS-routing problem and has to involve tight routing/MAC cross-layer design. For example, variation of bandwidth demand on a given routing path or change of a routing path can trigger reallocation of time slots, channels, antenna directions, etc., on all links related to the given routing path or *vice versa*.

Based on the above analysis, we know that cross-layer design is imperative for WMNs.

III. Cross-Layer-Design Protocols and Optimization Algorithms

A. General Methodology of Cross-Layer Design

Cross-layer design can significantly improve the network performance [6]–[8]. It can be performed in two ways: loosely coupled and tightly coupled cross-layer design.

In the loosely coupled cross-layer design, optimization is carried out without crossing layers but focusing on one protocol layer. In order to improve the performance of this protocol layer, parameters in other protocol layers are taken into account. Thus, information in one layer must be passed to another layer. Typically, parameters in the lower protocol layers are reported to higher layers. For example, the packet-loss rate in the MAC layer or channel condition in the physical layer can be reported to the transport layer so that a TCP protocol is able to differentiate congestion from packet loss. As another example, the physical layer can report the link quality to a routing protocol as an additional performance metric for the routing algorithms.

It should be noted that information from multiple layers can be used on another layer to perform cross-layer design. With such information, there are two different methods in utilizing such information. The first one is the simplest case of cross-layer design, in which the information in other layers works just as one of the parameters needed by the algorithm in a protocol layer. The performance of this algorithm is improved because a better (more accurate or reliable) parameter is used, but the algorithm itself does not need a modification. For example, the physical layer can inform TCP layer of the channel quality so that TCP can differentiate real congestion from channel-quality degradation and, thus, can carry out congestion control more intelligently. In the second method, based on the information from other layers, the algorithms of a protocol have to be modified. For example, if a MAC protocol can provide a routing protocol about its performance, the routing can perform multipath routing to utilize spatial diversity. However, the change from a single-path routing to multipath routing needs a significant modification in the routing protocol rather than just a parameter adaptation.

In the tightly coupled cross-layer design, merely information sharing between layers is not enough. In this scheme, the algorithms in different layers are optimized altogether as one optimization problem. For example, for MAC and routing protocols in a multichannel time-division multiple-access (TDMA) WMN, time slots, channels, and routing path can be determined by one single algorithm. Due to optimization across layers, it can be expected that better performance improvement can be achieved by the tightly coupled cross-layer design than the loosely coupled scheme. However, the advantage of the loosely coupled design is that it does not totally abandon the transparency between protocol layers.

An extreme case of tightly coupled cross-layer design is to merge different protocol layers into one layer. According to the concept of "layering as optimization decomposition," this kind of design tries to improve the network performance by relayering the existing protocol stack. Merging multiple protocol layers into one layer keeps the advantage of tightly coupled cross-layer design. Furthermore, it can also eliminate the overhead in cross-layer information exchange. Interestingly, merging multiple protocol layers is not just a theoretical concept but has been seriously considered in real practice. For example, in the upcoming 802.11 standards for mesh networks, the routing protocol is being developed as one of the critical modules in the MAC layer. Such a merge between routing and MAC layers provides a great potential to carry out optimization between MAC and routing within the same protocol layer.

Cross-layer design can be realized between multiple layers or between just two layers. Given a protocol stack, cross-layer design can be based in any combination of two protocol layers.

In the following sections, instead of going through all combinations of cross-layer design, we will focus on the ones that are most critical for WMNs. Considering the TCP/IP protocol architecture, the protocol layers that contain most specific features of WMNs include MAC, routing, and physical layer. In some cases, the transport layer needs to be optimized with physical layer in WMNs. Thus, in the remaining part of this section, we will investigate the detailed protocols in cross-layer design between MAC and physical, between MAC and routing, and between physical and transport layers. Optimization algorithms across multiple layers are also discussed.

B. MAC/Physical Cross-Layer Design

Cross-layer design between MAC and physical layers is more common than that between any other two layers, because MAC and physical layer are so close to each other. In many wireless networks, the lower part of the MAC layer and the baseband of the physical layer are implemented on the same card or even on the same chipset. Real-time interactions between the two layers occur frequently. Thus, in most wireless networks, including WMNs, the cross-layer between MAC and physical layer always exists in nature. On top of these natural interactions between the lower part of the MAC and the baseband of physical layer, the advanced physical-layer techniques have empowered the physical layer to be able to support more sophisticated cross-layer design for the purpose of enhancing network performance. These techniques include the following typical categories.

1) Multiple coding and modulation schemes. When a different coding and modulation scheme is used, the transmission rate on a link also changes.
2) Advanced antenna techniques. The examples include directional antennas and smart antennas.
3) MIMO. Based on multiple antennas for transmission and reception and advanced signal-processing techniques, the transmission rate of a wireless link can be significantly increased by MIMO.
4) Orthogonal frequency-division multiplexing (OFDM) technologies. OFDM can be used to build OFDM/TDD, OFDM/FDD, or OFDMA systems, as specified in IEEE 802.16. It can also be used as a building block for ultrawideband (UWB) systems.
5) UWB. Very high transmission rate is achieved using ultrawide bandwidth. UWB can be pulse-based like direct-sequence (DS) UWB as specified by UWB forum [9] or OFDM-based like multiband-OFDM (MB-OFDM) supported by WiMedia Alliance [10].

These technologies can be combined into one device. For example, a WiMedia UWB device, UWB is based on MB-OFDM, multirate is supported through variable coding and modulation, and link throughput can be improved through MIMO. The advanced physical-layer technologies provide a great potential of improved performance of delay, throughput, packet loss, etc. However, the physical layer itself does not determine how to adaptively fine tune the parameters in these advanced technologies. In fact, such fine tuning is a critical task in the upper sublayer of a MAC protocol. Thus, to optimize the performance of these advanced physical-layer technologies, the cross-layer design between MAC and physical layer becomes indispensable.

1) Adaptive Link Adaptation, Rate Control, and Framing: In a wireless network, fading, interference, noise, and so on can greatly impact the link capacity and, in turn, decreases the network capacity. To maintain a robust link performance, the most well-known technique is link adaptation through adaptive modulation and coding.

Link adaptation is coupled with rate control, because different modulation and coding schemes result in different transmission rate and also in different link-layer performance such as packet-error rate. Given a specific link, link adaptation dynamically selects the most appropriate modulation and coding scheme and, thus, the best transmission rate. Thus, as far as the transmission rate is concerned, link adaptation serves the same purpose as the rate control does. However, there exist some differences between rate control and link adaptation. In a rate-control scheme, the optimization is performed on the link transmission rate, while optimization is done directly on modulation and coding parameters in a link-adaptation scheme. Furthermore, link adaptation usually depends on physical-layer parameters such as bit-error rate (BER) or signal-to-noise ratio (SIR) to determine the coding and modulation parameters. Thus, the implementation of link adaptation is closer to the physical layer. One shortcoming of link adaptation is that the physical layer may lack a mechanism of providing accurate measurement of BER or SIR. On the other hand, in the MAC layer, the link-quality information can be derived from other easily measurable parameters, such as packet-loss rate, retransmission rates, etc., since such parameters change as the link quality varies. As a result, a different mechanism, i.e., rate control, is usually applied in the MAC layer to adaptively select the modulation and coding schemes in the physical layer. A rate-control scheme usually consists of two major modules: rate selection and mapping between rate and physical-layer parameters. Several MAC-layer parameters, such as packet-loss ratio, retransmission rates, packet-error rates, can be used as link-quality index to determine the best transmission rate. Given a transmission rate, the modulation and coding scheme can be selected based on a mapping table between rate and coding/modulation.

Most existing rate-control or link-adaptation schemes focus on link-layer performance. However, solely optimization on link rate or coding/modulation is not enough to guarantee the performance. For example, in either link adaptation or rate control, the link transmission rate needs to be reduced when BER or packet-loss rate increases. However, such a simple scheme may not be always working, because the BER or packet-loss rate may be just due to the inside-network interference between different nodes rather than the noise or outside-network interference. Thus, if a node's transmission rate is reduced, its transmission time is also increased and, thus, causes a higher duration of interference to other nodes in the same network. Other nodes performing the same rate-control or link-adaptation scheme experience the same problem, and then, the entire system becomes a positive-feedback close-loop control system, which means that the system can quickly loose stability and the rate in all nodes becomes very low. To solve this issue, the adaptive frame size in the MAC layer must be determined by considering the interference between different nodes in the same network. Such an adaptive framing scheme is a more advanced rate-control mechanism, which not only selects the best transmission rate but also determines the most appropriate frame size corresponding to this rate.

An example of rate-adaptive framing is proposed in [11], where the size of a MAC-layer frame is determined by a receiver and then fed back to the transmitter. Such a scheme can significantly achieve much better performance than the other rate-control schemes in [12] and [13].

It should be noted that simple link-adaptation or rate-control schemes are commonly used in the many current WMNs. For example, in IEEE 802.11-, 802.15-, and 802.16-based WMNs, all existing rate-control schemes are still based on rate-control schemes with optimization on either rate or modulation/coding only. However, as a multihop mesh network, the interactions between different nodes significantly impact the performance of the rate-control schemes. Thus, it is highly desired that the schemes, like rate-adaptive framing in [11], be developed for WMNs.

2) Adaptive Antenna-Direction Control: Compared to omnidirectional antenna, directional antennas hold several advantages. With a directional antenna, the same transmit power on a node can make signals reach much longer distance. In other words, for the same distance, a directional antenna can achieve much better link quality than an antenna with omnidirection. A directional antenna can effectively reduce the number of interfering nodes, which is particularly true in WMNs.

To take these advantages, the physical layer of a wireless node must be able to coordinate antenna directions in different nodes. Thus, the cross-layer optimization works as follows. First, the MAC determines the direction of a node. Second, the physical layer should be able to tune the antenna to the target direction.

In the physical layer, the simplest directional antenna is that the antenna is mechanically directional. However, such antenna is not scalable in WMNs, since the antenna direction of any node needs to be tuned to a different direction adaptively according to the variations of traffic pattern, link quality, network topology, etc. Another type of directional antenna is the sectored antenna. By using such an antenna, the antenna direction can be tuned to a certain sector. A more sophisticated way of achieving directional antenna is beamforming in a smart antenna. Given a target direction, the antenna beam can be formed to such a direction. Beamforming can achieve a more accurate antenna direction and have a finer granularity in tuning the directions.

In a wireless network with a point-to-point or point-to-multipoint topology, the adaptive antenna control is straightforward. However, when a WMN is concerned, the antenna-direction control becomes complicated, since a node may need to communicate with other nodes in different directions. Adaptive antenna-direction control reduces the exposed nodes in a WMN and, thus, has great potential to significantly increase the throughput. However, it also results in more hidden nodes. To avoid the performance degradation by these hidden nodes, the scheduling becomes a critical task. The simple scheme such as RTS/CTS mechanism defined in 802.11 is not effective anymore, because the hidden nodes are not due to the distance but are due to uncoordinated change of antenna directions on different nodes. Thus, the scheduling scheme does not reside on one node. Instead, it resides on different nodes in WMNs and runs as a distributed but cooperative algorithm among all nodes in WMNs. Since antenna change by the scheduling scheme also impacts the routing path, adaptive antenna-direction control actually involves a cross-layer design among three layers, i.e., routing, MAC, and physical layers.

3) Dynamic Subcarrier Allocation and Frame Aggregation for OFDM: OFDM has been used in many existing wireless networks including IEEE 802.11 and 802.16. In many OFDM-based wireless networks, the subcarriers in one OFDM symbol are treated as one resource unit. For example, in 802.16 wireless networks, the TDMA/FDD and TDMA/TDD modes do not allow subcarrier allocation. In 802.11 networks, subcarrier is not visible to the MAC-layer protocol. However, as the physical-layer transmission rate becomes higher and higher, the subcarrier allocation becomes necessary. Considering a TDMA frame in which a time slot contains one OFDM symbol, if the physical-layer transmission rate is high, then one time slot can hold a large packet size. To avoid underutilization of an OFDM symbol, frame aggregation is needed in the MAC layer. However, the effectiveness of frame aggregation depends on enough traffic load generated on a node. In addition, frame aggregation causes performance degradation such as the increased latency of a packet. To avoid such issues, subcarrier allocation is preferred. With subcarrier allocation, finer resource unit can be achieved, and thus, a packet can be accommodated as it arrives.

There is another motivation for subcarrier allocation. In a wireless network, particularly in a multihop wireless network, nodes in the same network can experience different fading. As a result, for the same subcarrier, it may experience bad channel quality on one node but good channel quality on another node. Thus, it is beneficial to allocate different subcarriers to different nodes whose channel quality varies depending on their locations.

In IEEE 802.16, the operation mode OFDMA provides an option of subcarrier allocation. In Qualcomm's Flash OFDM, subcarrier allocation is also supported. In a point-to-multipoint wireless network, subcarrier allocation has been thoroughly researched [14]–[16]. However, in a multihop wireless network, such as WMNs, few results have been reported on subcarrier allocation. In [17], subcarrier allocation is studied for a WMN with a single mesh router and multiple mesh clients. A two-layer hierarchical fair-scheduling scheme is proposed to determine subcarriers and their powers for the mesh router and mesh clients. Slow fading is considered in the proposed scheduling scheme. Thus, such a scheme is not applicable to frequency-selective fading channels.

4) MIMO Control and Scheduling: It is well known that MIMO can significantly improve the link capacity of a wireless network via transmit diversity and spatial multiplexing. Such a technique has been considered as the most important solution to extend physical transmission rate of IEEE 802.11 wireless networks, i.e., in the upcoming 802.11n standards. MIMO in a wireless network can be treated independently from the MAC protocol. This method of protocol design is simple but is definitely a suboptimal solution; the advantageous features of MIMO seen in a point-to-point link may not be maintained in a more complicated network topology as in WMNs. To fully utilize the advantages of MIMO, the MAC protocol must be particularly designed, considering the cross-layer dependence with the MIMO physical-layer techniques.

By using multiple antennas for transmitting and receiving, several performance improvements can be achieved in a MIMO system [18].

1) Transmit diversity: The same information is sent on different antennas to increase the reliability, which, in turn, increases the throughput in the MAC layer.
2) Spatial multiplexing: Different streams of packets are sent on different antennas and, thus, achieve a higher transmission rate than a single-antenna system.
3) Beamforming: Different transmission angles are controlled in different antennas so that a desired beam is formed pointing to a certain direction. With beamforming, better transmission range and higher rate can be achieved.
4) Interference nulling: Again, via control on different antennas, the interference from or to certain directions can be reduced. Thus, the interference between different nodes can be controlled.

The above improvements are not mutually exclusive and can be combined to reach even higher performance improvement. How to trigger different functionalities in the physical layer so as to get the best combination of the earlier improvements is one of the tasks of cross-layer design between MAC and physical layers. In a WMN, in addition to the earlier improvements, another improvement is also possible.

Since each node has multiple antennas and multiple neighboring nodes to send and receive packets, it can send packets to different nodes using different antennas or it can receive packets from different nodes using different antennas. Such multiuser OFDM puts a more challenging requirement on scheduling of packet transmissions in a MIMO system.

MIMO control and scheduling usually consist of the following critical steps.

1) Get channel-state information (CSI). CSI can be obtained at the receiver from the training sequence in a receive signal. However, there exist three difficulties of getting CSI. First, the training data is sent before MIMO control takes effect; otherwise, it is impossible for a node to get CSI for all neighbors. However, the dilemma is that the CSI can be much different from the case when MIMO control kicks in. Thus, identifying the right set of reachable neighbors and the CSI from them with MIMO transmissions is still an unresolved and challenging research issue. Second, the CSI feedback to transmitter may not always be available, for two reasons. One reason is that the channel may change so quickly that the feedback is too slow to catch up the change. The other reason is that there is a lack of a mechanism of sending the feedback information from the receiver to the transmitter. Without CSI at the transmitter, an open-loop system needs to be developed to carry out MIMO control. Third, the mobility or topology change makes the previous two problems even more severe since neighbors of a given node are not stable, and thus, the CSI from these neighbors must be updated constantly.

2) Determine the tradeoff between transmission rate, range, and reliability. This is an independent step from the previous one. Usually, the network performance metrics such as throughput and QoS are important factors to determine the needed MIMO control, such as spatial multiplexing, diversity, beamforming, or interference nulling. The challenge in this step is how to combine as more features as possible so that the best performance can be achieved.
3) Scheduling packet transmissions on different antennas on different nodes using the collected CSI. Based on the collected CSI and determined MIMO control, scheduling schemes need to be developed for packet transmissions on different nodes. In WMNs, the challenge is on how to develop a distributed scheduling scheme such that the global optimal solution can be obtained.

A few research results have been reported to investigate the optimization of MIMO transmissions [19], [20]. However, these results are derived based on simple assumptions such as perfect synchronization in packet transmission [19] and no physical-layer dependencies and channel variations [20]. Furthermore, same as directional antenna control, MIMO control and scheduling may also involve routing protocol as part of cross-layer design. So far, no research has been reported on this topic.

C. Routing/MAC Cross-Layer Design

A routing protocol of a multihop wireless network determines a path for any packet from its source to destination. In its simplest form, a routing protocol can just consider connectivity between nodes, i.e., as long as connectivity can be maintained, a routing path is set up. However, to enhance performance, other routing metrics and mechanisms must be taken into account. For example, a routing protocol may need to consider minimum hop count, lowest traffic load, etc. However, such types of layered design approaches are still suboptimal in performance. The reason is that the behavior of the MAC protocol has not been taken into account. Thus, no matter how the routing protocol is optimized, if the underlying MAC does not provide satisfying performance, then the overall performance perceived by a routing protocol can be poor.

A MAC protocol aims to provide medium-access opportunities to nodes sharing the same medium, given any condition of traffic load, interference, noise, and topology of a network. However, traffic load, interference, and so on are closely related to a routing protocol. Thus, the performance of a MAC protocol can be significantly impacted by a routing protocol.

In order to achieve the best network performance, routing and MAC must be jointly optimized.

1) Methodology of Routing/MAC Cross-Layer Design: Routing/MAC cross-layer can be done in a simple loosely coupled scheme as follows. A routing protocol collects information in the MAC layer, such as link-quality, interference-level, or traffic-load information, to determine the best routing path. Such a method can only achieve a limited performance gain, since the MAC layer is considered but not optimized accordingly.

In order to optimize the performance of routing and MAC protocols together, the working mechanisms of a MAC protocol must be explored and optimized as part of the tasks of routing/MAC cross-layer design.

It is well known that a MAC protocol can be reservation or random-access based. For a random-access-based MAC, no mechanism is available to fine-tune the MAC layer performance by considering information from the upper layer. Instead, a node just tries its best to access the medium. Such a MAC has a great advantage of simplicity and has another advantage of being decoupled from upper protocol layers. However, the shortcoming is that the MAC itself has low performance, and routing protocol can even have worse performance since no chance of cross-layer optimization is available. Such a problem reflects one of the many issues of applying CSMA/CA MAC protocol to WMNs. There are two possible solutions to this problem. One is to modify the random-access protocol so that it becomes closer to a reservation protocol. For example, the 802.11e hybrid channel-access control includes mechanism of scheduling and reservation, which works together with CSMA/CA to improve the performance of 802.11 MAC. The other solution is to have overlay protocols. For example, we can develop a TDMA protocol overlaying CSMA/CA [21].

Due to limited capabilities of MAC/routing joint optimization for a random-access network, we start to focus on cross-layer design between a routing protocol and a reservation-based MAC protocol. Although today's WMNs are still mostly based on CSMA/CA-type random-access MAC, more and more WMNs are starting to use reservation-based MAC. One reason is that many existing multihop wireless networks are being standardized under the framework of TDMA. Typical examples include 802.16 mesh networks and relaying networks, UWB mesh networks, Wimedia mesh networks, etc. Another reason is that the poor performance of CSMA/CA for WMNs have motivated the development of better MAC protocols overlaying CSMA/CA. With such an enhancement, the overall MAC works approximately as a reservation-based MAC.

A reservation-based MAC protocol is usually concerned with scheduling packet transmissions with respect to properly assigned resources. Thus, the critical task in such a MAC is resource-allocation-considering constraints such as QoS, interference, network topology, etc., which are all related to a routing protocol. The network resources can be time slots, code-division multiple-access (CDMA) codes, channels, and so on. In order to have optimal resource allocation, the routing path and resource allocation can be determined in the same algorithm. This algorithm can be split into suboptimization problems into both MAC and routing layers or can be merged into one protocol layer: either MAC or routing.

Joint optimization between time-slot (or code) allocation and routing has been reported in work on QoS routing in multihop wireless networks [22], [23]. In WMNs, a router is usually powerful enough to operate using multiple channels or even multiple radios. Such a capability adds new dimension for resource allocation, and thus, new joint optimization scheme is needed between channel allocation and routing path.

2) Joint Channel Allocation and Routing: Channel allocation depends on how traffic is distributed in the network, which is determined by routing. However, given the same routing paths, different channel allocation will also result in different network performance. Thus, joint optimization between channel allocation and routing protocol is an important topic for WMNs. As of the present, some research papers reported some solutions for this optimization problem. While most of them are focused on the optimization algorithm [24], [25], some other papers have proposed practical protocols [26].

a) Joint channel allocation and routing protocol: In [26], a joint channel allocation and routing scheme, called Hyacinth, is proposed for WMNs. It is assumed that the traffic aggregated at mesh routers only goes to or comes from the gateway nodes. With such an assumption, spanning trees are built up from gateway nodes to all nongateway nodes. For each end-to-end traffic flow, the routing path is found along spanning trees such that the load balancing is considered. After a routing path is set up, channels are assigned to NICs on each node via a local channel-allocation scheme. Thus, the following major functions exist in [26].

1) **Load-balancing routing**. The key step of load-balancing routing is to construct the routing tree. When a tree is being built up, the routing metric takes into account traffic load along the tree. Thus, a routing path based on such a tree will achieve load balancing. In [26], the following routing metrics are considered.
   a) Hop count. This is number of hops from a node to the wired network.
   b) Gateway link capacity. This is the residual capacity of the uplink that connects the root gateway of a tree to the wired network.
   c) Path capacity. This is the minimum residual capacity on the path from a node to the wired network.

   The gateway link capacity and path capacity are dynamic depending on the interactions of multiple nodes. Thus, route flaps can occur and result in nonconvergent network behavior. In order to avoid such an issue, when a new node joins a tree, a "join" message should propagate all the way back to the gateway node to check if the join is acceptable. If not, the new node cannot join the tree. Although such a scheme can avoid route flaps, but the whole process is complicated and slow and causes more overhead to the routing protocol.
2) **Distributed load-aware channel assignment**. Given a routing path, the channels in all nodes along this path need to be assigned. In any multiradio network, channel allocation in a node usually impacts channel allocation on other nodes, which is called the channel-dependency issue. In order to resolve such an issue, two sets of NICs are specified for each node: UP-NICs and DOWN-NICs. Channel allocation at a node is only performed on DOWN-NICs, and UP-NICs use the same channels as those in DOWN-NICs of its parent node.

   In order to give more bandwidth to nodes closer to the root of a tree and to not impact the channel allocation in parent nodes, higher priority is given to parent nodes when channel assignment is performed. Thus, when a node searches for a channel, it is restricted to those channels that are not used by interfering neighbors with a higher priority.

Simulation study shows that the Hyacinth protocol can significantly improve the throughput. Experiments based on a prototyping system also illustrate that the Hyacinth protocol is fast in routing-path recovery. However, the joint routing and channel assignment scheme still contains the following shortcomings.

1) The Hyacinth protocol totally depends on spanning trees. The validity of such a scheme depends on an assumption that the traffic of all nodes goes or comes to/from the gateway nodes. For other traffic patterns, the protocol does not work anymore.
2) The channel-dependency problem still exists in all children nodes and nodes in the same level at the same level. When a node's DOWN-NICs are updated with new channels, the channel allocation in all its children nodes and nodes at the same level of the spanning tree also have to be updated.
3) Channel allocation may not be convergent. Channel allocation can be started by any nodes due to the distributed method. Although priority is given to parent nodes, priority between nodes in the level on the spanning trees are the same. Due to uncoordinated allocation of nodes with the same priority, their channel assignment may not be convergent and, thus, may cause severe interference among nodes.
4) Channel assignment and load-balancing routing may be inconsistent. Channel assignment is performed based on a routing path established considering load balancing. However, when channels are reassigned for different nodes on the routing path, the actual load along this routing path, as well as the capacity in other links, will also be changed. Thus, although load-balancing routing and channel assignment are decoupled in the Hyacinth protocol, it does not mean that the two functions will produce consistent results. In other words, when channel assignment is done, the routing path may have lost the advantage of load balancing.
5) Traffic-load estimation does not necessarily reflect the actual traffic load. This is because the MAC-layer contention cannot be accurately captured by weighted summation.

3) Advanced Features and Challenges: In a single-channel WMN, resource allocation can be done as that in the framework of QoS routing. It should be noted that the resource at the MAC layer is not fixed. It can be variable due to the variance of channel quality and changing parameters in the physical layer. Such variations result in fluctuations of link capacity, which is usually regulated by a rate-control algorithm. Thus, a WMN is usually a multirate system.

Since the transmission rate is not just related to link quality, it also impacts the transmission range (and, thus, the topology), interference, etc. As a result, rate control is coupled with both resource allocation and routing. When multicast is considered, this problem becomes even more challenging [27]. How to carry out joint optimization among resource allocation, rate control, and routing is still an open research topic.

When multichannel operation is also considered, the earlier problem becomes even more sophisticated. Thus, joint optimization among channel allocation, rate control, resource allocation, and routing becomes one of the most difficult problems for routing/MAC cross-layer design.

D. Transport/Physical Cross-Layer Design

1) Motivations and Approaches: In a multihop wireless network, the capacity of a link is usually variable due to factors such as interference, time-varying channel quality, fading, and so on. Without a fixed capacity in these links, an end-to-end transmission mechanism, i.e., a transport-layer protocol, needs to be optimized by considering the varying link capacity. This motivates the need for cross-layer design between transport-layer protocol and physical-layer techniques.

Transport-layer protocol can be simple or complicated, depending on what services need to be provided at the transport layer. The two most well-known transport-layer protocols are TCP and UDP. For UDP, the mechanism is very straightforward; a source node just sends its desired traffic rate without considering what will happen in the intermediate nodes and links from itself to the destination node. TCP works significantly differently. A source node needs to adaptively adjust its transmission rate according to the congestion condition in the network. The congestion can be real congestion on a certain link or poor quality in a link.

Because of different transport mechanisms in TCP and UDP, their impact to the overall performance of the network is quite different. UDP does not obey any rule of controlling traffic rate at the transport layer. Thus, in order to improve the overall network performance, the source rate must be regulated by other mechanisms such as connection admission control or end-to-end rate control. Due to the variable link capacity, these control algorithms must be cross-optimized with the physical layer. Thus, the cross-layer design between UDP and the physical layer becomes a problem of joint optimization between physical layer and admission control or rate control.

On the other hand, for a TCP protocol, a congestion-control algorithm must exist to regulate the source rate. Thus, cross-layer design between TCP and physical layer is a problem of joint optimization between the congestion-control algorithm of TCP and different physical-layer parameters.

In the remaining part of this section, we focus on the scenario of TCP and physical-layer joint optimization.

2) Cross-Layer Optimization Between TCP and Physical Layers: Cross-layer design between TCP and physical layer for a multihop wireless network has been researched for several years. Different methods in the literature can be classified into two categories. In the first category, the congestion-control algorithm of TCP is optimized by considering the information collected from the physical layer. One example is to use the physical-layer information to differentiate packet loss due to congestion from that due to link-quality-related loss. Such optimization can only achieve limited performance improvement, because the interaction between TCP and physical layer is not considered. However, when a link is congested, the physical layer can adjust its parameters, e.g., transmit power, to avoid congestion, which will also help TCP achieve better performance. Similarly, when a link experiences low quality, the physical-layer parameters, such as coding rate or transmit power, can be adjusted to enhance the link quality. Thus, instead of passively taking action only in TCP, TCP and physical-layer-control schemes can be jointly optimized. Such schemes belong to the second category of cross-layer design between the TCP and the physical layer. Different from the first category, the second category involves more complicated algorithms as well as more sophisticated protocols and their implementations. Because of such challenges, many research issues remain unresolved.

Congestion-control mechanisms have been analyzed as distributed algorithms that solve the NUM problem [28], [29]. For physical layer, as more advanced technologies are developed, its control becomes more and more complicated, particularly when cross-layer design is involved. Thus, the key tasks of joint optimization between congestion control and physical layer is twofold: One is to extend the existing congestion-control-optimization algorithm to embrace the physical-layer factors, and the other is to determine what parameters need to be controlled in the physical layer as well as to optimize such parameters together with congestion control.

There exist many variants of TCP, such as Tahoe, Reno, Vegas, etc. However, the congestion-control mechanisms of all of them follow the same rule: The transmission rate of each source is adjusted based on implicit or explicit feedback of congestion signals generated by active queue management. Some of them use loss as congestion signals, while others use delay. Since delay-based congestion signal pertains good properties of convergence, stability, and fairness [28], it is favored by many existing congestion-control algorithms. TCP Vegas is one of the congestion-control algorithms that use the delay-based congestion signal. If d_s is the propagation delay from a source to its destination and D_s is the delay of both propagation and congestion-induced queuing delay, then the TCP window w_s needs to be updated by considering the difference between these two parameters; more specifically, the difference of the expected rates w_s/d_s and w_s/D_s. D_s is measured based on the timing information in acknowledgment (ACK) packets. Thus, the sliding window of TCP Vegas can be updated as follows:

$$w_s(t+1) = \begin{cases} w_s(t) + \frac{1}{D_s(t)}, & \text{if } \frac{w_s(t)}{d_s} - \frac{w_s(t)}{D_s(t)} < \alpha_s \\ w_s(t) - \frac{1}{D_s(t)}, & \text{if } \frac{w_s(t)}{d_s} - \frac{w_s(t)}{D_s(t)} > \alpha_s \\ w_s(t), & \text{otherwise} \end{cases} \quad (2)$$

where α_s is a parameter that controls the congestion level and impacts the stable transmission rate.

The physical layer has many parameters to be controlled. The most well-known ones include transmit power, coding, and modulation. Other parameters include antenna direction, beamforms, etc. Thus, it is difficult to have one control mechanism that covers the optimization of all such parameters. A more practical scheme is to focus on one or two parameters in the control mechanism and assume that the others are fixed. For example, in [30], power control is considered as the main mechanism of fine-tuning the physical-layer performance. As follows, we discuss how joint optimization between TCP congestion control and physical-layer power control can be done [30].

a) Joint TCP congestion control and power control: The joint TCP congestion control and power control can be formulated as a problem of optimizing users' utility with respect to the transmit power and users' source rate.

For a source s, suppose its rate is x_s and utility is $U(x_s)$. Thus, for all sources, the overall utility is $\sum_s U_s(x_s)$.

Looking at a link l, its capacity c_l is determined by the transmit power of itself, noise, and other users. Suppose that the power levels of all nodes are denoted by a vector $\mathbf{P}$, then we have [30]

$$c_l(\mathbf{P}) = \frac{1}{T}\log(1 + K\text{SINR}_l(\mathbf{P})) \tag{3}$$

where T is the symbol period, K is a constant depending on BER and modulation, and SINR_l is the signal-to-interference-noise ratio. Usually, $K = -(\phi_1/\log(\phi_2(\text{BER})))$, where ϕ_1, ϕ_2 are constants depending on modulation. Considering a CDMA-like system and denoting G_{lk} as the path gain from the transmitter of link l to the receiver of link k, then SINR_l can be described as $(P_l G_{ll}/\sum_{k\neq l} P_k G_{lk} + n_l)$, where n_l is the noise at the receiver of link l. Given a link l, all traffic passing through it cannot exceed its capacity $c_l\mathbf{P}$. If the set of links on the routing path of source s to its destination is $L(s)$, then link l is subject to a constraint of $\sum_{s:l\in L(s)} x_s \leq c_l(\mathbf{P})$.

As a result, joint congestion control and power control needs to find an optimal solution of rate $\mathbf{x} = x_1, x_2, \ldots$ and power $\mathbf{P}$ such that the overall utility is maximized, i.e., the cross-layer optimization can be formulated as follows [30]:

$$\begin{aligned} &\text{maximize} && \sum_s U_s(x_s) \\ &\text{subject to} && \sum_{\substack{s:l\in L(s)\\ (\mathbf{x},\mathbf{P})\geq 0}} x_s \leq c_l(\mathbf{P})\forall l. \end{aligned} \tag{4}$$

As shown in this optimization problem, the optimization of congestion is performed not only over source rate but also over transmit power level, which is much different from Internet congestion-control algorithm. The earlier joint optimization scheme makes several assumptions in the network protocol. First, the physical and MAC layers work as a CDMA system. Second, the routing path is assumed to be single path and is predetermined. Last, the coding rate and modulation types are also assumed to be fixed.

To derive a concrete solution, delay-based congestion signal and TCP Vegas sliding window update are considered in [30]. More specifically, the utility function can be further described as $U_s(x_s) = \alpha_s d_s \log(x_s)$, where α_s is a parameter used in TCP Vegas sliding window update and the window-update procedure described in (2) is followed. The detailed results of the joint optimization problem in (4) are given in [30] and provide two cross-related iterative equations to update sliding window, transmission rate, and transmit power.

b) Discussion on the algorithm: The earlier cross-layer optimization algorithm can be implemented in a distributed way. However, a node needs to flood certain information like SINR, power level, and path gain to all other nodes. In addition, the measurement of such information may not always be accurate due to stochastic characteristics of SINR and path gain. Thus, the robustness of the earlier algorithm to the fluctuations of these measured parameters needs to be studied. It has been proved in [30] that the proposed joint congestion-control and power-control algorithm is robust to parameter perturbation, and the convergence can be achieved in a geometric rate. It also has a nice property of global convergence to the optimal solution $(\mathbf{x}^*, \mathbf{P}^*)$. Graceful tradeoff between algorithm complexity and performance improvement can be also achieved.

However, the joint congestion-control and power-control optimization algorithm is limited to the scenario where several assumptions must be satisfied.

First of all, coding and modulation is fixed; otherwise, the optimization algorithm needs to determine the optimal selection of coding rates and modulation schemes. For some multihop wireless network, in particular, some low-rate networks, this assumption is reasonable. However, for WMNs that are usually concerned with high-speed transmission, the physical layer is always expected to adaptively adjust coding rate and modulation. Such work is usually performed in a rate-control algorithm in the MAC layer. However, because of the change of rate, the parameters such as transmit power and link capacity also change. Thus, the joint optimization between TCP and physical layer needs to consider variable coding and modulation.

The function modeling the relationship between link capacity and power control may be different in many WMNs. For CDMA-based WMNs, this model works fine. However, many WMNs are not based on CDMA but on TDMA or random access. For these WMNs, SIR is very small when multiple nodes send packets simultaneously. Although the capture effect helps some nodes receive correct packets even under interference, an interfered node usually cannot send or receive correct packets. Thus, fine-tuning power does not have significant impact to the link capacity. In other words, joint power control and congestion control will not achieve optimal throughput performance. For such WMNs, the more critical task is to carry out joint optimization between congestion control and scheduling in the MAC layer, as will be discussed in Section III-E.

In the joint congestion- and power-control algorithm, the routing path is assumed to be fixed. In fact, when congestion occurs, another well-known mechanism is to find a better routing path so that congestion is avoided. Such a mechanism can be achieved through multipath routing or load-balancing routing. This proves that cross-layer design between transport layer and physical layer is not enough and can inevitably involve routing protocol. Such issues will also be discussed in Section III-E, but we note that, so far, no effective solutions are available to provide joint optimization across all protocol layers.

E. Joint Optimization Algorithms Across Multiple-Protocol Layers

For a multihop wireless network like WMN, the design of the entire protocol stack can be formulated as one optimization problem. We call this approach as "full-optimization design." A solution to this problem can be mapped to different protocol layers in the clean-slate protocol architecture. Such an approach can achieve a layered design or loosely coupled cross-layer design, since the interactions between layers are small due to optimization throughout the protocol stack. However, the shortcoming is that the protocol layers derived from optimization may not exactly match an existing protocol stack like the Internet. In order to avoid such an issue, another approach is to formulate an optimization problem considering the existing protocol architecture. Thus, it is just a "suboptimization design." A solution to this problem provides no help in reducing the interactions between protocol layers but can significantl y improve the performance by optimizing cross-layer interactions.

Cross-layer optimization can be formulated across different protocol layers ranging from the application to the physical layer. For example, the cross-layer design in Section III-D illustrates the cross-layer optimization between physical and transport layer. However, in WMNs or a multihop wireless network, in general, the most typical cross-layer optimization is the joint optimization of congestion control and scheduling. In a multihop wireless network, congestion control can be done end-to-end in transport layer or link-by-link in the MAC layer, while scheduling involves the close interactions between MAC and physical layer. A scheduling algorithm determines the parameters for both MAC and physical layers and depends on the congestion control to determine a best transmission rate. The interactions between the congestion control and the scheduling also involve the routing protocol. Thus, a well-defined joint optimization between congestion control and scheduling can enhance performance optimization in layers, including transport, routing, MAC, and physical layers.

1) Joint Optimization of Congestion Control and Scheduling: A network user usually expects to get as more resources as possible from the network. When multiple users are considered and if no arbitration is enforced, then they can easily end up in a situation where none of them can really get satisfactory service quality. A conventional solution to this problem is to use transport-layer protocol to perform congestion control, routing protocol to find the best path considering load balancing, and MAC protocol to schedule transmissions with an objective to achieve the best one-hop performance. However, such a scheme cannot achieve optimal performance, because the algorithms in different protocol layers are not optimized altogether. In other words, the network is not really optimized with an objective to satisfy as more users as possible.

To improve the network performance, joint optimization between transport, routing, MAC, and physical layer is needed. Such an optimization should be performed with an aim of maximizing users' interests. Since the transport is mostly concerned with congestion control and MAC/physical layer is concerned with scheduling, algorithms on joint congestion control and scheduling provide a promising approach to optimize network performance to maximize the benefits of networks users.

In a joint congestion control and scheduling algorithm, transport-layer protocol is considered in the congestion-control part, MAC and physical layers are considered in the scheduling part, and routing is embedded in the interactions between congestion control and scheduling. The optimization target of such an algorithm is to maximize the users' benefits as defined by utility functions. With this in mind, we will discuss how the algorithm of joint congestion control and congestion is formulated in the next section.

a) Formulations: To formulate joint optimization between congestion control and scheduling, we need to define two models that capture the behaviors of congestion control and scheduling. The objective of congestion control is to find each user's rate such that the utilities of all users are maximized under the condition that the network system can be stabilized by a certain scheduling scheme. Thus, the objective of scheduling is to design a scheduling policy such that the given rate of each user is satisfied in a stable system.

Assume that there are K users in the network. Given a user k, its traffic originates from source node s_k and d_k has a rate of r_k. We assume that the rate r_k is upper bounded by M_k. The utility of user k is a function of rate, i.e., it is $U_k(r_k)$. Thus, the congestion control and scheduling must achieve the following joint optimization:

$$\max_{r_k < M_k} \sum_k U_k(r_k)$$
$$\text{subject to} \quad \vec{r} \in \Lambda \tag{5}$$

where $\vec{r} = [r_k, k = 1, \ldots, K]$ is a vector of all users' rates. Λ stands for the rate region or called the capacity region that contains the set of all rate vectors for which a scheduling scheme can be found to stabilize the network. Thus, the congestion control and scheduling are cross-related via this rate region. For example, given an optimized rate vector, what scheduling scheme can be used is also constrained. On the other hand, the set of all available scheduling schemes also constrains the best rate vector in (5). It should be noted that, in (5), the overall utility of the entire network is considered via summation of all users' utilities, which is just a simple but reasonable scenario. However, more sophisticated methods of integrating different users' utility functions can also be considered.

To solve the optimization problem in (5), the first step is to derive the capacity region Λ. Two schemes can be used to define the capacity region: node-centric and link-centric capacity region [32]–[35].

b) Node-centric capacity region and optimal solution: Considering a network with N nodes, links between these nodes are represented by the set $\mathcal{L} = \{(i,j), i,j = 1,\ldots,N, i \neq j,\}$, and the capacity of each link (i,j) is c_{ij}. In order to ensure users' rate falls into a stable capacity region, the incoming and outgoing traffic rate on each node must be balanced. Based on such a concept, the node-centric capacity region can be derived. Denote c_{ij}^d as the link capacity used by traffic toward destination d on link (i,j). Thus, a user-rate vector belongs to the capacity region if and only if the following constraint is satisfied [32], [33]:

$$\sum_{j:(i,j)\in\mathcal{L}} c_{ij}^d - \sum_{j:(j,i)\in\mathcal{L}} c_{ji}^d \geq \sum_{k:s_k=i,d_k=d} r_k$$
$$\text{for all } i \text{ and all } d, \text{ and } i \neq d,$$
$$r_{ij}^d \geq 0 \text{ for all } (i,j) \in \mathcal{L} \text{ and for all } d. \tag{6}$$

Considering that the capacity set of all links is $\mathcal{C}$, since the capacity of all links $\vec{c}$ cannot lie outside the convex hull of $\mathcal{C}$, a second constraint for the capacity region needs to be considered

$$\left[\sum_d c_{ij}^d \in Co\left((\mathcal{C})\right)\right]. \tag{7}$$

To further consider the impact of the physical layer, the capacity of each link is related to a few other factors such as power control, modulation, coding, rate control, and so on. Thus, the capacity of all links is $\vec{c} = \mathcal{F}(\vec{P})$, where $\vec{P}$ is a vector representing all physical-layer parameters. If the set of all feasible physical-layer parameters is Π, then the capacity set of all links is $\mathcal{C} = \{\mathcal{F}(\vec{P}), \vec{P} \in \Pi\}$. Although the function $\mathcal{F}$ is usually nonconvex, the convex hull of $\mathcal{C}$ is convex and is also closed and bounded. Thus, the capacity region Λ is a convex set. The optimization problem of (5) can be solved with the help of Lagrange multiplier q_i^d for the constraint of each (i, d) in (6). Thus, the joint-congestion-control-and-scheduling problem is decomposed into the following two cross-related subproblems: congestion control for calculating the data rate of each user and scheduling for the derivation of the transmission rate of each link. The detailed solutions to these two subproblems are presented in [33] and [34]. From these results, it is necessary to note that multipath routing is implicitly assumed in the above joint optimization between congestion control and scheduling.

c) Link-centric capacity region and optimal solution: Link-centric capacity region can be derived by considering the balanced traffic load on each link. In order to know how and who contributes the traffic on a link, a routing matrix has to be specified for each user on each link. If the traffic of user k passes through link (i, j), then the routing index for this user on this link, denoted by H_k^{ij}, is one; otherwise, $H_k^{ij} = 0$. Thus, the user-rate vector $\vec{r}$ lies in the capacity region Λ if and only if

$$\sum_{k=1}^{K} H_k^{ij} r_k \leq c_{ij} \text{ for all } (i, j) \in \mathcal{L}, \vec{c} \in Co(\mathcal{C}). \tag{8}$$

To consider the physical-layer impact on the link capacity, the same relationship between $\vec{c}$ and physical-layer parameters $\vec{P}$ as that for node-centric capacity region can be used. The solution to the joint-congestion-control-and-scheduling-optimization problem can be derived in a similar way as that for the node-centric case [35], i.e., joint congestion control and scheduling is decomposed into the following two crossly related components: congestion control and scheduling.

d) Comparisons: There exist several key differences between node-centric and link-centric joint congestion control and scheduling optimization [35].

The first difference lies in how routing is handled in the joint-optimization problem. In the node-centric scheme, the routing protocol, assumed to be multipath routing, is considered in the scheduling component. In the link-centric scheme, routing path is predetermined, so no routing protocol is actually considered.

The traffic model reflects the second difference. In the link-centric scheme, the balance equation has an implicit assumption that the traffic on different links on the routing path is the same. This is only true when a user's traffic is constant or there is no delay in delivering traffic. However, there is no such assumption in the node-centric scheme. In this sense, the node-centric scheme has a more accurate traffic model.

In addition to the differences, both schemes share some similarities. Both schemes assume a centralized optimization algorithm. However, when we come across a wireless multihop network like WMN, a distributed scheme is needed. How to map these two schemes into a distributed scheme is not self-explanatory in the derivations and needs further research. In addition, both schemes can take into account various physical-layer characteristics such as channel variations in the scheduling component.

2) Limitations of Cross-Layer Optimization Algorithms:

a) Perfect versus imperfect scheduling: In the scheduling component of the joint-congestion-control-and-scheduling-optimization problem, the optimal solution may be difficult to derive. For example, the Lagrange multiplier used in the optimization changes every time period, which implies that the scheduling must be updated per time period. This results in a highly complicated scheduling scheme and renders the optimization algorithm nearly useless in practical implementation. In order to lower the complexity, the following two approaches can be taken.

In the first approach, the optimization algorithm is only applied to a simple network model. For example, in an infrastructure network like a typical wireless LAN setup. An optimal schedule can be achieved with polynomial-time complexity. For a node-exclusive interference model where only two nodes can communicate at the same time, an optimal schedule can also be achieved with low complexity. An example of this case is the Bluetooth network. However, this approach is not applicable to WMNs, since the network model is much different from that of WMNs.

In the second approach, we have to relax the optimality requirements of the scheduling problem. Although scheduling with such relaxed requirements becomes imperfect, the capacity region that can be achieved is much smaller than that of perfect scheduling. As a result, the complexity of the scheduling scheme is much lower. Imperfect scheduling and its impact to cross-layer optimization are studied in [36].

b) Implementation issues: Besides complexity, the cross-layer optimization algorithms also have several other critical issues.

When coming to the applicability of cross-layer optimization algorithms, the first question we have to face is how to map the algorithms onto the existing protocol stack. For example, WMNs are usually built based on the well-known Internet protocol architecture, in which a variant of TCP protocol is applied to control the network congestion, the MAC protocol varies as different physical-layer technologies are used, and different types of routing protocol can be used. Unless we use a totally clean-slate protocol architecture, we have to modify formulation of the cross-layer optimization algorithm.

In the joint congestion control and scheduling algorithm, the MAC layer is assumed to be schedulable. In fact, some MAC-layer protocols are totally random. A typical example is CSMA/CA protocol, which is widely accepted as the basic MAC for IEEE 802.11 wireless networks. For such random MAC protocols, scheduling the transmission rate for each link cannot be easily achieved. Thus, to develop stochastic scheduling scheme so that the optimal rate of each link can be achieved is a challenging problem. Another approach is to design another better coordinated MAC, overlaying the random MAC [21], [37], and then, the optimal scheduling is performed on top of the overlaying MAC protocol.

Between MAC and physical layer, another difficult problem is how to accurately model the relationship function between MAC and physical layers. This is even more challenging for WMNs, since the physical layer usually contains many sophisticated technologies such as MIMO, adaptive coding/modulation, adaptive power control, multichannel operation, and so on.

The existing cross-layer optimization schemes have not taken into account the QoS requirements by users. In the previous section, the joint congestion control and scheduling algorithms only achieve the optimal rate for each user. However, a user usually does not care about what user rate he shall obey to achieve the best performance of the entire network. Instead, what he is really interested in is QoS expectation. For example, he may demand a traffic rate that is not an optimal solution of the cross-layer optimization but has to be satisfied in order to meet his QoS requirement. In this case, the cross-layer optimization problem needs to have a new formulation.

As shown in the joint congestion control and scheduling algorithms, the routing protocol is not considered in a proper way. More specifically, the connectivity of a routing path is not ensured in these algorithms. No matter how a routing protocol is assumed (either multipath routing or predetermined routing), we need to make sure that the routing path for a given user is connected; otherwise, such a routing mechanism makes no sense. In other words, the optimal solutions derived on the basis of such a routing protocol is not useful in practical implementations.

In view of the earlier issues, we believe cross-layer optimization will continue to be a challenging research topic for WMNs.

IV. Concluding Remarks

As explained in previous sections, there is no doubt that the cross-layer design can definitely improve the network performance. However, issues can come together with benefits, as explained as follows.

1) System complexity: For many cross-layer design schemes, it can easily be shown that they achieve great performance through simulations or even prototypes. However, when coming to the actual implementation of these schemes, we face to several complexities in modifying protocols in different layers. These modifications can impact the maintainability of the software, stability of different protocol modules, and flexibility of porting codes to different platforms.
2) Protocol interoperability and compatibility: With cross-layer design, the standard working mechanism in the protocol stack is broken. Thus, a wireless network with cross-layer design can easily be incompatible with other networks, and thus, interoperation between different networks is difficult to maintain. Consequently, a cross-layer design scheme should have a remedy to standard compatibility. However, it can be imagined that, even if interoperation can be maintained via such a remedy, the benefits of cross-layer design may diminish when networks with and without cross-layer design have to work together.
3) Evolution capability: In a layered-protocol architecture, protocols in one layer can evolve separately without disrupting the functionalities of protocols in another layer. When cross-layer design is adopted, any upgrade or change in protocols must be coordinated among different protocol layers. This requirement significantly limits the capability of product evolution.

It should be noted that such issues usually do not exist in a layered design scheme. To avoid such issues, tradeoff should be made between performance improvement via cross-layer design and benefit loss of layered design. However, technically, it is extremely difficult to carry out a reasonable tradeoff since issues such as system complexity or protocol interoperability is not easy to quantitatively be evaluated. Thus, in this paper, we suggest several rules that can be followed to avoid blind use of cross-layer design.

1) Achieve enough margin of performance improvement. Cross-layer design brings network-performance improvement with a price of high system complexity. Thus, to compensate the cost, the performance improvement must be significant enough. Multiple performance metrics may need to be considered together to evaluate the overall network-performance improvement. In fact, using cross-layer design, we can easily see some performance improvement in throughput, delay, packet loss, etc. However, if the improvement is only a small percent, e.g., 5%, then it is not a wise strategy to adopt cross-layer design, since such performance improvement can easily vanish due to uncertainties in a wireless network like interference, noise, shadowing, etc.
2) Explore any possible opportunity that can improve network performance using layered-protocol design. For cross-layer design, benefits always come together with issues. Thus, the best strategy is to explore the capability of layered-protocol design as much as possible. The theoretical research work on "layering as decomposition optimization" can be used as guidelines in doing so.
3) Carry out cross-layer design without compromising framework specified by standards. In order to ensure standard compatibility to the great extent and, thus, to maintain interoperability and evolution capability, it is a good strategy to carry out cross-layer design under the framework of standard specifications.
4) Push standardization of cross-layer design framework and methodology. To further improve the viability of cross-layer design schemes, standardizing the framework of cross-layer design is necessary.

Moreover, several principles discussed in [3] can be adopted as additional cautionary guidelines for cross-layer design.

1) Take into account the dependency graph for the entire protocol stack. With cross-layer design, protocols become interactive with each other. The interactions can be dependent on each other and cause multiple adaptation loops, which causes performance degradation in protocols not being considered in the cross-layer design. To solve this problem, a dependency graph representing interactions between protocols needs to be derived for the entire protocol stack.
2) Time-scale separation and stability. Based on the dependency graph, if a parameter is controlled by two different adaptation loops, time-scale separation can be used to avoid conflict. The rationale is that the two entities controlling the same parameter work on different time scale. Adaptive control theory has proved that the stability can be achieved via time-scale separation. If they work in similar time scale, then the close-loop control theory should be used to prove the stability of the given interactions.
3) Avoid unbridled cross-layer design. If multiple cross-layer interactions are employed, then it is easy to get an unstructured spaghetti-like protocol architecture, which is hard to maintain. In addition, in this case, network-performance improvement may be only achievable within a small area of equilibrium state. Away from equilibrium, the network performance can be much worse than what can be achieved by a layered-protocol design.

These principles can help avoid unintended consequences of using cross-layer design.

REFERENCES

[1] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Commun. Mag.*, vol. 43, no. 9, pp. S23–S30, Sep. 2005.

[2] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proc. IEEE*, vol. 95, no. 1, pp. 255–312, Jan. 2007.

[3] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *Wireless Commun.*, vol. 12, no. 1, pp. 3–11, Feb. 2005.

[4] H. Jiang, W. Zhuang, and X. Shen, "Cross-layer design for resource allocation in 3G wireless networks and beyond," *IEEE Commun. Mag.*, vol. 43, no. 12, pp. 120–126, Dec. 2005.

[5] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[6] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: Joint routing, scheduling, and power control," in *Proc. IEEE Annu. Conf. INFOCOM*, 2004, pp. 1457–1466.

[7] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "A framework for cross-layer design of energy-efficient communication with QoS provisioning in multi-hop wireless networks," in *Proc. IEEE Annu. Conf. INFOCOM*, 2004, pp. 1446–1456.

[8] M. Chiang, "To layer or not to layer: Balancing transport and physical layers in wireless multihop networks," in *Proc. IEEE Annu. Conf. INFOCOM*, 2004, pp. 2525–2536.

[9] [Online]. Available: http://www.uwbforum.org

[10] [Online]. Available: http://www.wimedia.org

[11] C.-C. Chen, H. Luo, E. Seo, N. H. Vaidya, and X. Wang, "Rate-adaptive framing for interfered wireless networks," in *Proc. IEEE Annu. Conf. INFOCOM*, 2007, pp. 1325–1333.

[12] G. Holland, N. H. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," in *Proc. ACM Annu. Int. Conf. MOBICOM*, 2001, pp. 236–251.

[13] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proc. ACM Annu. Int. Conf. MOBICOM*, 2002, pp. 24–35.

[14] C. Y. Wong, R. S. Cheng, K. B. Letaief, and D. Ross, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[15] X. Wang and W. Xiang, "An OFDM-TDMA/SA MAC protocol for broadband wireless LANs with QoS constraints," *ACM/Kluwer Wireless Netw.*, vol. 12, no. 2, pp. 159–170, Jan. 2006.

[16] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks—Part I: Theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614–624, Mar. 2005.

[17] K.-D. Lee and V. C. M. Leung, "Fair allocation of subcarrier and power in an OFDMA wireless mesh network," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp. 2051–2060, Nov. 2006.

[18] M. Zorzi, J. Zeidler, A. Anderson, B. Rao, J. Proakis, L. A. Swindlehurst, M. Hensen, and S. Krishnamurthy, "Cross-layer issues in MAC protocol design for MIMO Ad Hoc networks," *Wireless Commun.*, vol. 13, no. 4, pp. 62–76, Aug. 2006.

[19] P. Casari, M. Levorato, and M. Zorzi, "On the implications of layered space-time multiuser detection on the design of MAC protocols for ad hoc networks," in *Proc. IEEE PIMRC*, 2005, pp. 1354–1360.

[20] K. Sundaresan, R. Sivakumar, M. A. Ingram, and T.-Y. Chang, "A fair medium access control protocol for ad-hoc networks with MIMO links," in *Proc. IEEE Annu. Conf. INFOCOM*, 2004, pp. 2559–2570.

[21] X. Wang, W. Wang, and M. Nova, "Distributed TDMA for wireless mesh networks," Mar. 9, 2005. U.S. Patent Application 11/076738.

[22] T.-W. Chen, J. T. Tsai, and M. Gerla, "QoS routing performance in multihop, multimedia, wireless networks," in *Proc. IEEE ICUPC*, 1997, pp. 557–561.

[23] C. R. Lin and J.-S. Liu, "QoS routing in ad hoc wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1426–1438, Aug. 1999.

[24] M. Alicherry, R. Bhatia, and L. Li, "Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks," in *Proc. ACM Annu. Int. Conf. MOBICOM*, 2005, pp. 58–72.

[25] J. Tang, G. Xue, and W. Zhang, "Interference-aware topology control and QoS routing in multi-channel wireless mesh networks," in *Proc. 6th ACM Int. Symp. MOBIHOC*, 2005, pp. 68–77.

[26] A. Raniwala and T.-C. Chiueh, "Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network," in *Proc. IEEE Annu. Conf. INFOCOM*, 2005, pp. 2223–2234.

[27] W. Ge, J. Zhang, and S. Shen, "A cross-layer design approach to multicast in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1063–1071, Mar. 2007.

[28] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Control Syst. Mag.*, vol. 22, no. 1, pp. 28–43, Feb. 2002.

[29] S. H. Low, "A duality model of TCP and queue management algorithms," *IEEE/ACM Trans. Netw.*, vol. 11, no. 4, pp. 525–536, Aug. 2003.

[30] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 104–116, Jan. 2005.

[31] S. H. Low, L. L. Perterson, and L. Wang, "Understanding Vegas: A duality model," *J. ACM*, vol. 49, no. 2, pp. 207–235, Mar. 2002.

[32] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936–1948, Dec. 1992.

[33] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," in *Proc. IEEE Annu. Conf. INFOCOM*, 2003, pp. 745–755.

[34] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," in *Proc. IEEE Conf. Decision Control*, 2004, pp. 1484–1489.

[35] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.

[36] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer congestion control in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 302–315, Apr. 2006.

[37] X. Wang, V. Rollinger, A. Patil, C. Gui, W. Wang, and M. Nova, "Distributed multichannel wireless communications," May 26, 2006. U.S. Patent Application 11/420668.

Ian F. Akyildiz (M'86–SM'89–F'96) received the B.S., M.S., and Ph.D. degrees from the University of Erlangen-Nurnberg, Erlangen, Germany, in 1978, 1981, and 1984, respectively, all in computer engineering.

He is currently the Ken Byers Distinguished Chair Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, and the Director of the Broadband and Wireless Networking Laboratory. His current research interests are in cognitive radio wireless networks, wireless sensor networks, and wireless mesh networks.

Prof. Akyildiz is an Editor-in-Chief of the *Computer Networks Journal* (Elsevier) as well as the founding Editor-in-Chief of the *Ad Hoc Network Journal* (Elsevier) and the *Physical Communication Journal* (Elsevier). He was the recipient of the "Don Federico Santa Maria Medal" for his services to the Universidad of Federico Santa Maria, Valparaiso, Chile, in 1986. From 1989 to 1998, he served as a National Lecturer for the Association for Computing Machinery (ACM), and he was the recipient of the ACM Outstanding Distinguished Lecturer Award in 1994. He was also the recipient of the 1997 IEEE Leonard G. Abraham Prize Award (IEEE Communications Society) for his paper entitled "Multimedia Group Synchronization Protocols for Integrated Services Architectures" published in the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS in January 1996 and was the recipient of the 2002 IEEE Harry M. Goode Memorial Award (IEEE Computer Society) with the citation "for significant and pioneering contributions to advanced architectures and protocols for wireless and satellite networking." He received the 2003 IEEE Best Tutorial Award (IEEE Communication Society) for his paper entitled "A Survey on Sensor Networks," which was published in the *IEEE Communications Magazine* in August 2002. He was also the recipient of the 2003 ACM Sigmobile Outstanding Contribution Award with the citation "for pioneering contributions in the area of mobility and resource management for wireless communication networks." He was also the recipient of the 2004 Georgia Tech Faculty Research Author Award for his "outstanding record of publication of papers between 1999 and 2003." He was also the recipient of the 2005 Distinguished Faculty Achievement Award from the School of Electrical and Computer Engineering, Georgia Institute of Technology. He has been a Fellow of ACM since 1996.

Xudong Wang (M'03) received the B.E. degree in electrical engineering and the Ph.D. degree in automatic control from Shanghai Jiao Tong University, Shanghai, China, in 1992 and 1997, respectively, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, in August 2003.

He is currently the R&D Manager with Kiyon, Inc., San Diego, CA, where he leads the research and development of wireless mesh networking technologies. He was also a Senior Network Architect and Researcher with Kiyon, Inc. He is the creator of Kiyon's pioneering single-channel and multiple-channel time-division multiple-access technologies for wireless mesh networks. His research interests also include cognitive/software radios, cross-layer design and communication protocols for cellular networks, mobile *ad hoc* networks, sensor networks, and ultrawideband networks. He is the holder of a number of patents on wireless-networking technologies, and most of his inventions have been successfully transferred to products.

Dr. Wang is an Editor for *Elsevier Ad Hoc Networks* and was also a Guest Editor for several journals. He was the Demo Cochair of the Association for Computing Machinery (ACM) International Symposium on Mobile *Ad Hoc* Networking and Computing (ACM MOBIHOC 2006). He is currently a Technical Program Cochair of the Wireless Internet Conference (WICON) 2007. He has been a technical committee member of many international conferences and a technical reviewer for numerous international journals and conferences. He is a member of ACM and ACM SIGMOBILE and a voting member of the IEEE 802.11 and 802.15 Standard Committees.