Cross-Layer Framework for QoS Support in Wireless Multimedia Sensor Networks

Ghalib A. Shah, *Member, IEEE*, Weifa Liang, *Senior Member, IEEE*, and Ozgur B. Akan, *Senior Member, IEEE*

***Abstract*—The emergence of wireless multimedia sensor networks (WMSNs) has made it possible to realize multimedia delivery on tiny sensing devices. The volume and characteristics of multimedia data is quite different from the data generated in WSNs that has raised the need to explore communication protocols for multimedia delivery in WMSNs. The existing studies focus on providing quality-of-service (QoS) to each individual source but they are not adaptive to create room for maximizing the number of sources. In this paper, we propose a novel cross-layer framework for QoS support in WMSNs. The objective of the proposed framework is to maximize the capacity of the deployed network to enhance the number of video sources given that the QoS constraint of each individual source is also preserved. This is achieved by implementing Wyner-Ziv lossy distributed source coding at the sensor node with variable group of pictures (GOP) size, exploiting multipath routing for real-time delivery and link adaptation to enhance the bandwidth under the given bit error rate. Hence, application requirements are mapped on joint operations of application, network, link and MAC layers to achieve the desired QoS. Simulation results reveal that the framework admits larger number of video sources under the satisfied distortion constraint.**

***Index Terms*—Cross-layer QoS, IEEE 80211e, wireless multimedia sensor networks (WMSNs).**

I. Introduction

WIRELESS multimedia sensor networks (WMSNs) are rapidly gaining interest of researchers due to the great success of wireless sensor networks (WSNs) in solving real-world problems, such as environmental monitoring, health monitoring, security surveillance, precision agriculture, and industrial processes automation. They are generally composed of wirelessly interconnected small devices that allow fetching video, audio and still images and reporting to a sink node [1]. Recently, the research in WMSNs gained high momentum due to the introduction of the new vision of Web of Things [32], [33], which is evolving very rapidly considering the tremendous progress in the field of embedded systems that has given birth to a great deal of tiny embedded devices and also the never ending progress towards sophisticated communication techniques. This has given a new dimension to WSNs and its derivatives, WMSNs, to explore the possibility of getting recognition of their segregated efforts by the rest of networking community. Moreover, the support of low-power devices 6LowPAN [34] in the development of IPv6 also inspires this study since it gives an opportunity to integrate WMSNs with the Internet enabling remote surveillance and monitoring.

In traditional multimedia systems, a streaming server takes the media input from the special-purpose sophisticated devices and broadcasts it over wideband links to a number of users (one-to-many) subscribed for the service. The server is capable of performing all the computationally intensive operations to encode the media contents using any well-known encoding technique, e.g., MPEG. The decoder on the other hand is a relatively lightweight device operating in a slave mode to the encoder. In contrast, multimedia sensors in WMSNs have severe resource constraints including bandwidth, energy, storage, computation, and, therefore, complex encoding techniques are not feasible to be applied. The communication paradigm in WMSNs is also many-to-one in which a sink is the only ultimate destination of all the sensor nodes and is capable of performing complex decoding operations. Moreover, data produced by the sensors exhibit highly spatial correlation due to the high density of node deployment.

Multimedia support in various wireless networks, such as broadband wireless networks, satellite networks, mobile ad hoc networks, has been widely explored [16], [18], [31]. However, there are several main peculiarities that make multimedia content delivery in wireless sensor networks challenging, which is largely unexplored. Most of these challenges are unique due to the limited resources imposed on sensor nodes, and hostile environment resulting in variable channel capacity. Transporting video with the guaranteed quality-of-service (QoS) in WMSNs is of prime importance due to higher rate requirements on limited and variable capacity channels. Layered approaches may achieve very high performance in terms of the metrics related to each of these individual layers. However, they are not jointly optimized to maximize the overall network performance. Because of the direct coupling between the physical layer and the upper layers, the traditional protocol stack is not efficient in WMSN. By taking into account the resources in WMSNs, joint optimization of networking layers, i.e., cross-layer design, stands as the most promising alternative to inefficient traditional layered protocol architectures [1], [3]. A number of cross-layer [3], [5], [15], [28] protocols are proposed for QoS data delivery, which

Manuscript received January 04, 2012; revised March 01, 2012; accepted April 03, 2012. Date of publication April 25, 2012; date of current version September 12, 2012. This work was supported by the Turkish Scientific and Technical Research Council (TUBITAK) Career Award under grant #110E249. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Monica Aguilar.

G. A. Shah and O. B. Akan are with the Next-Generation and Wireless Communication Laboratory (NWCL), Department of Electrical and Electronics Engineering, Koc University Sariyer, Istanbul 34450, Turkey (e-mail: gshah@ku.edu.tr; akan@ku.edu.tr).

W. Liang is with the School of Computer Science, Australian National University, Canberra, Australia (e-mail: wliang@cs.anu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2012.2196510

mainly focus on providing QoS to each individual stream according to the current network state, while they are not adaptive to make room for the new sources if the available bandwidth is not sufficient rejecting the admission straightforward.

In this paper, we investigate the cross-layer design for maximizing the number of video sources such that the quality of each video stream is also preserved. More specifically, it is a bandwidth enhancement problem subject to the distortion perceived at the sink, bandwidth reservation for each individual source, real-time packet delivery and packet reliability constraints. Sensor nodes implement conventional encoding algorithm for key frame generation and Wyner-Ziv lossy distributed video coding for smaller size frames, where the ratio of key and distributed coded frames is maintained as a tradeoff between distortion and source bit rate. These frames are routed using a novel source directed multipath routing (SDMR) algorithm. SDMR interacts with the underlying MAC protocol and maintains three disjoint directed paths from source to destination, which are selected according to bandwidth, delay and reliability requirements. MAC functions are based on QoS supporting IEEE 802.11e MAC standard with some enhancements in scheduling to address fairness and power saving. At the link layer, we consider multirate transmission modes of IEEE 802.11 and a selection of a particular mode is made dynamically according to bandwidth requirement with constrained link error rate. Hence, application requirements are mapped on joint operations of application, network, MAC and link layers to achieve the desired QoS.

The remainder of the paper is organized as follows. Section II provides an overview of the existing approaches considering QoS at individual layers or through cross-layer signalling in WMSNs. In Section III, system model is presented with formal problem definition. The proposed QoS mechanism exploiting cross-layer signalling is presented in Section IV. Detailed performance evaluation and simulation results are discussed in Section V. Finally, in Section VI we draw the main conclusion.

II. Related Work

In this section, we investigate the existing studies, which provide complete or partial QoS support for wireless multimedia sensor networks. These studies can be grouped into cross-layer approaches and layered approaches with routing and MAC as two important candidate layers.

A. Cross-Layer Approaches

Cross-layer design has been advocated as the state-of-the-art solution for resource constrained wireless networks [1] and a great deal of research is being conducted on this perspective. Melodia *et al.* proposed a cross-layer QoS architecture for WMSNs in [3], which is based on hop-to-hop QoS contract, receiver-based traffic scheduling, Time-Hopping Impulse radio based MAC and dynamic channel coding algorithm. QoS parameters are negotiated from the source to the destination prior to the flow admission using handshaking. There are a number of limitations in this architecture. First, it relies on the reserved path and ignores the higher failure rate of sensor nodes that in turn causes higher path failures. Second, hop-to-hop four-way handshaking is an energy exhaustive operation in WMSNs and thus incurs formidable overheads. Third, traffic scheduling is based on TDMA, which in turn depends on the time synchronization of nodes. Incorporating the characteristics of the data or application level information to cross-layer designs assists a great deal in achieving optimized data delivery. In [15], the entropy tracking algorithm is proposed to adapt the transmission rates of sensor nodes based on environmental variations and underlying correlation structure. The correlation in this algorithm is merely based on the distance between nodes, which does not reflect the spatial correlation in video sensors since there is a specific view range of these devices for capturing videos or images. A transport, network and physical cross-layer solution is provided in [5] that operates in two phases. It aims to maximize data gathering while minimize end-to-end delay. Transport layer estimates the end-to-end delay and only those routes meeting the end-to-end delay constraints are selected, while the physical layer adjusts the transmission radius and rate in order to meet delay deadlines as well as data rate in a given expected network lifetime. End-to-end delay based route selection cannot provide persistent QoS since the change in network conditions will be slowly reflected in adapting routes.

Although the objective of [28] is similar to the proposed framework, however, the solution is quite different in many ways. It does not consider energy efficiency, and no admission control and link adaptation is implemented. The approaches in [29] and [30] partially address the problem of providing QoS support to maximum number of multimedia sources and only seek for adaptive routing and link that can provide more capacity. However, they do not consider the energy constraints of sensor nodes in routing apart from lack of admission control. Moreover, they do not implement the classification of traffic and prioritizing the video traffic according to the importance factor of each individual frame. Nevertheless, the existing cross-layer approaches do not exploit the multipath routing potential to address the high bandwidth demands and also there is no control to preserve the QoS for each multimedia flow when the large number of flows originate.

B. Routing Layer

The key challenge for routing protocols in WMSNs is to optimize network performance for required QoS support and energy conservation. Geographic routing protocols [9], [10] are more suitable for WMSNs since they choose the shortest path to the destination efficiently. However, these protocols need to exploit the possibility of multiple paths to enhance the available bandwidth for high rate multimedia delivery. EQSR [10] is a multipath routing protocol that provides service differentiation to allow delay sensitive traffic to meet deadlines, end-to-end delay guarantees as well as reliability through redundancy. Similarly, a geographical directional routing (GDR) protocol [9] is proposed for video sensor networks, which is based on directed disjoint multipaths from the source to the sink. This approach is suitable when there are only few video streams otherwise it may result in flooding. Moreover, duplicate packets may cause congestion and in turn consume excessive energy for transmission of extra packets including the routing cost. The authors of [13] proposed a multipath based multiconstrained routing protocol

for WSN. It exploits greedy routing for path selection subject to reliability and delay constraints. Similarly, MMSPEED [11] adopted a routing-MAC cross layer approach for real-time delivery of packets through multiple paths. These studies do not consider the bandwidth requirement that is critical for video sensors and, thus, are not suitable for WMSNs. In [12], a greedy routing algorithm based on directed diffusion is proposed. Its objective is to maximize throughput and minimize delay. However, this protocols does not exploit the multiple paths to provide higher bandwidth, which is a preliminary requirement for WMSNs. Hence, the existing routing protocols do not address the issues of WMSNs in a unified way.

C. MAC Layer

Although ZigBee or IEEE 802.15.4 standard is devised for low power devices to operate at short range, it is inevitable to be employed for WMSN due to its low rate of 50–250 kbps. Similarly, ultra wideband (UWB) is also evolving to support high rate for low power devices but they have very short transmission range and are inapplicable to WMSNs requiring range in order of few tens of meters to more. Whereas IEEE 802.11 standard has already been widely accepted as a physical layer technology and medium access control for mobile ad hoc networks and it is commonly employed in many commercial-off-the-shelf video devices, which are largely deployed for video surveillance and monitoring applications. Moreover, its amendments [2] for QoS support has made it a potential MAC candidate for WMSNs with little modification for energy conservation. In the literature, there exist many MAC layer schemes providing differentiated access in traditional wireless sensor networks, by varying inter frame size and contention window size [6]–[8]. A QoS-aware MAC protocol [6] for WMSN is proposed that addresses the bandwidth and delay constraint in addition to energy conservation through differentiated duty cycling. In this protocol, video packets and best effort traffic are assigned different priorities, which are mapped to multiple traffic classes of different contention window sizes. PQ-MAC [8] is an interesting scheme that provides service differentiation through a doubling strategy. It adds an advanced listening slot during its sleep cycle just to forward high priority data. Thus it aims to reduce latency of delay sensitive traffic. MAQ [7] integrates CSMA/CA and TDMA to support fair and predictable medium access for delay sensitive traffic. It provides interleaving of CFP and CP to allow low latency channel access to delay sensitive traffic. Falah *et al.* [31] has proposed an efficient transmission of H.264 video streams using IEEE 802.11e. The scheme provides service guarantee by mapping the video streams to controlled channel access phases (CAPs) in the contention-free period and pre-empting the contention period if in progress. Such an idea is useful only for infrastructure based wireless networks and cannot be adopted for WMSNs. These approaches mainly lacks in providing the fairness to multiple streaming flows and does not adopt the link to the multimedia distortion perceived at the receiver.

Our Contribution: There exists no unified solution for efficient multimedia delivery of large number of video streams in WMSNs. This paper addresses the implementation of maximizing the number of video sources such that the quality of individual source does not get affected. The existing studies focus on guaranteed QoS through end-to-end resource negotiation and setting up routing paths. In contrast, we adopt stateless routing mechanism and resource negotiation is localized within the neighborhood to avoid from the high overhead. Moreover, we incorporate application level information to optimize the communication layers, which is not explored yet.

III. System Model

A WMSN consists of n stationary video sensor nodes which are randomly dense deployed and a sink node (base station) that is used to monitor a region of interest. We assume that all the nodes as well as the sink are location-aware and the location of sink is known to all the sensor nodes. Moreover, the transmission range (r) of each sensor node is identical and each is equipped with omnidirectional antenna. If the sensors are deployed in the deployment area A, then the average node density $\rho(r)$ in transmission radius r is computed by $\pi r^2/An$. For the sake of simplicity, we use ρ to represent the node density in a circular region of radius r in the rest of this paper.

A. Application Model

Due to higher node density, video streams are spatially correlated that result in data redundancy. We use distributed source coding (DSC) [17] at application layer to overcome data redundancy, in which the spatially correlated sources encode their video by assuming that the sink has already acquired side information to jointly decode. In our source coding model, source nodes encode VIDEO with mixed encoding consisting of conventional video encoder to generate key frames, referred to as **I**-Frames, for side information computation and lossy Wyner-Ziv encoder exploiting side information to generate much smaller frames, referred to as **W**-Frames. A group of pictures (GOP) of size g comprises of an **I**-Frame and $g-1$ **W**-Frames. For tolerable encoding distortion δ_e, the output rate of the source i is computed by the following formula [26]:

$$r_i(\delta_e) = \frac{1}{g}\left[r^I(\delta_e) + \sum_{i=1}^{g-1} r^{WZ}(\delta_e)\right] \tag{1}$$

where $r^I(\delta_e)$ and $r^{WZ}(\delta_e)$ are the bit rates of **I**-Frame and **W**-Frame of source i, respectively. In a dense video sensor network, the multiview image coding [27] has been used to exploit spatial correlation of video sensor nodes by employing geometric prediction model that reduces the decoding losses. Such a scheme only requires application level changes at the sink and does not effect the cross-layer framework.

B. Network Model

We further assume that the routing algorithm is a unicast and adopts to multiple paths to manage the *bandwidth*, *delay*, and *reliability* constraints (which will be discussed in Section IV-A). The routing algorithm is also a stateless in which the selection of a forwarding node is purely based on the localized decision of the relaying node and thus, fixed routing paths for the requests are not established yet. We divide the forwarding region into three zones for the three non-interfering (NI) paths through each zone as shown in Fig. 1. Let N_1^i, N_2^i, and N_3^i be the subset

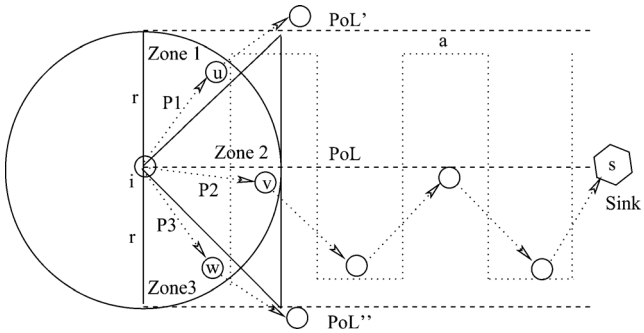

Fig. 1. Multipath forwarding zones to establish non-interfering paths.

of neighboring nodes of i in its three zones. Three paths P_1, P_2, and P_3 are established corresponding to each zone, which are originated at links u, v, and w destined to the sink, where $u \in N_1^i$, $v \in N_2^i$, and $w \in N_3^i$.

Let $R_{in}^i(t)$ and $R_{out}^i(t)$ represent the total incoming data rate from the upstream nodes of i and the total outgoing data rate from i to its downstream nodes through the paths P_1, P_2, and P_3, respectively, at any time instant t. The values are measured in a discrete time fashion in a sequence $\{0, 1, \ldots, t, t+1, \ldots\}$ of sampling intervals of length Δ_t. In a stateless routing protocol that does not establish path *a priori*, we assume that if per hop delay can be guaranteed under constrained number of hops, then end-to-end packet deadline is presumably met [11].

Now, let τ be the end-to-end packet deadline from any source to the sink. We use the minimum advancement (a) as *m/hop* towards the sink and determine the expected hops count $E[h_{ix}]$ from source i to sink x. By having hops count, we can compute the per hop delay budget $\tau^i = \tau / E[h_{ix}]$ for source i.

Furthermore, the MAC layer is based on enhanced distributed channel access (EDCA) of IEEE 802.11e [2], which is an enhancement for QoS in the legacy distributed coordination function (DCF) of IEEE 802.11. The multi-rate transmission modes of IEEE 802.11 are exploited in order to dynamically achieve the desired capacity and reliability. Assuming that the interface supports m number of transmission modes, $\chi_i(k)$ and $b_i(k)$ represents the bit error rate (BER) and the bandwidth of the interface operating at k^{th} mode selected by node i, respectively, for $0 \leq k < m$. Let Δ_t be the sampling period during which a node measures per hop delay, link reliability and available bandwidth. Moreover, a source sends its *request* only once in a sampling interval and if the request is rejected, then it attempts again in the next interval.

C. Distortion Model

The quality of decoded video is affected by three factors: encoding error (δ_e) due to quantization, transmission errors (δ_l) due to packet losses, and decoding error (δ_d) introduced due to the disparity of side information. However, the proposed framework aims to control the distortion due to packet losses and side information disparity due to variable group of picture size, which are essentially independent to each other. Hence, the overall distortion, D, is computed as sum of δ_l and δ_d[23]. The quality of video is assumed to be acceptable only if it is no greater than the tolerable distortion threshold (Θ), i.e., $D \leq \Theta$. When a packet does not arrive at the receiver by its playout deadline, the decoder conceals the missing information to avoid interruptions and the playout continues at the cost of higher distortion. The mean distortion introduced due to packet loss rate is modeled as

$$\delta_l = \sigma_{u_o}^2 P_L \sum_{i=1}^{g-1} \frac{1-\beta i}{1+\gamma i} \tag{2}$$

where $\sigma_{u_o}^2$ is a constant value that describes the sensitivity of the video decoder to an increase in error rate. If the decoder implements some advance error concealment techniques, then the value is kept correspondingly low. β and γ are two model parameters, where $\beta = 1/(g-1)$ and γ shows the efficiency of loop filter for removing error whose value can be obtained from some empirical results as discussed in [23]. Similarly, higher motion-compensation disparity or inaccuracy of the side information function produces decoding errors. In practice, the larger the value of g, the higher the motion-compensation disparity between the side information and distributed coded **W**-Frame is predicted. The distortion introduced due to motion-compensated interpolation at the decoder using Wyner-Ziv distributed coding is derived in [26]

$$\delta_d = \frac{1}{4\pi^2} \int \int_{\Lambda} \min \left[\theta, \hat{\phi_{ee}}(\Lambda)\right] d\Lambda \tag{3}$$

where θ is the rate-distortion curve control parameter, which is proportional to the error produced due to quantization and $\hat{\phi_{ee}}(\Lambda)$ is the spectral power density of motion-compensated prediction error computed in [26].

D. Problem Definition

Let $\Re$ be the set of video sensors sending their data to the sink in the network with meeting their constrained distortions, where $0 \leq |\Re| < n$. A new request can be accommodated if each relaying node has adequate available bandwidth and also meets the constraints for desired quality of video. If there is unlimited bandwidth, every new coming request can be fulfilled, which, however, is impossible in practice. To make the new request admissible, one strategy is to increase the current link capacity to a certain amount by using some link adaptation function in multi-rate wireless interface, such as IEEE 802.11, and adaptive video encoding. Consider a video stream request sourced at i as a sequence of **I**-Frames and **W**-Frames for a GOP size g and encoded data rate r_i. The source always keeps bandwidth reserved for its own video stream and its admission is controlled at the relaying nodes. If the request of source i was rejected by the relaying node j at time $t-1$ due to bandwidth limitation, node i increases the value of g to reduce its data rate and sends its request again at time t. Given a set of new requesting sources in the network, the problem is to maximize the admission of new requests at the relaying nodes. We define $\tilde{g}$ as the new size of GOP to reduce the source data rate $(\tilde{g} > g)$, $r_i(\tilde{g})$ as data rate of source i at $\tilde{g}$, $\bar{\tau_i}$ as the mean packet delay measured by node i for packet size l and P_L as the tolerable packet loss rate threshold.

For a new request of i, we formulate if this request is admissible at the relaying node j as

$$\textbf{maximize} \sum_{i\in\Re_\delta} (r_i(t-1) - r_i(t)) + b_j(k) - \left(R^j_{in}(t) + r_j(t)\right),\ j \in \cup_{y=1}^{3} N^i_y,\ 0 \le k < m-1 \quad (4)$$

subject to

$$D_i \le \Theta, \quad \forall i \in \Re \quad (5)$$

$$\left(R^j_{in}(t) + r_j(t) + r_i(t)\right) \le R^j_{out}(t), \quad i \in \Re_\delta,\ j \in \cup_{y=1}^{3} N^i_y \quad (6)$$

$$\bar{\tau_j}(t) \le \tau^i, \quad i \in \Re_\delta,\ j \in \cup_{y=1}^{3} N^i_y \quad (7)$$

$$\left(1-(1-\chi_{ij}(k))^l\right) \le \sqrt[E[h_{ix}]]{P_L}, \quad i \in \Re_\delta,\ 0 \le k < m. \quad (8)$$

In the objective function (5), we aim to accommodate the source i by reducing its data rate and increasing the link capacity of relaying node j using higher transmission mode. The first term is the gain in bandwidth acquired by reduction in source data rate in which source i had higher data rate at time $t-1$ and it is reduced at time t. At time $t-1$, the request might have been rejected previously and it attempts again at t with lower data rate. However, if it is the first attempt at time t, then r_i at $t-1$ is equal at t. The difference of the second and third terms gives the available bandwidth at the kth transmission mode. Constraint (5) ensures that the increase in GOP size does not cause to increase the distortion above the tolerable limit, while constraint (6) guarantees the availability of bandwidth to all admitted requests. Constraint (7) ensures the selection of a forwarding node whose mean packet delay is not higher than the per hop delay budget of the source. Constraint (8) makes sure that the selection of a new higher rate transmission mode does not cause to increase in packet losses above the tolerable packet loss rate. In the following section, we describe the distributed algorithm for the solution of the problems 5–9, while the global optimization of the problem could be searched in future work.

IV. Cross-layer QoS Framework

QoS for multimedia contents is generally characterized by four parameters; delay jitter, end-to-end deadline, bandwidth and reliability over unreliable medium. In multimedia applications, delay jitter is compensated by carefully choosing the playout time. However, packet deadline, bandwidth and reliability are dealt in the network at different communication layers jointly or separately. On the other hand, distortion is the metric to measure the perceived QoS at sink node. Yet, energy efficiency is the fundamental requirement for QoS support in WMSN that must need to be incorporated during all the network operations. In this framework, we consider per hop bandwidth (6), delay (7), and reliability (8) constraints such that the packets are delivered under probabilistic hop count limit as discussed in Section IV-A. These constraints are met collectively through interaction of application, routing, link and MAC layers in a cross-layer architecture as shown in Fig. 2. Whereas, the video distortion is estimated at the sink node

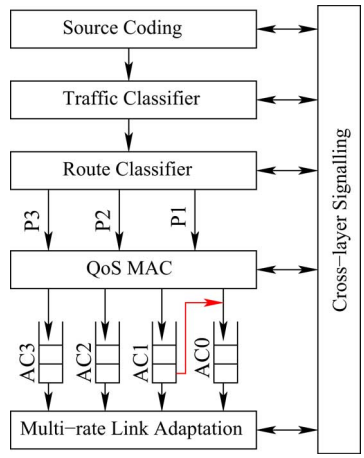

Fig. 2. Cross-layer QoS protocol architecture.

that sends feedback to the sources. Traffic classifier module classifies the types of frames, i.e., **I**-Frames and **W**-Frames, encapsulates the application data into a multimedia header having fields of *frame type*, *frame priority*, and *GOP size* and transfers it to the route classifier module. It is implemented over the routing layer and also coordinates with the application to adjust its rate, i.e., GOP size etc when the routing layer does not find sufficient capacity to route the data. Route classifier implements the proposed source directed multiple path routing (SDMR) algorithm and chooses an appropriate path for frame delivery. The link layer adaptation algorithm implemented at link layer dynamically chooses the appropriate transmission mode directed by the network layer to provide the required bandwidth subject to the link reliability. Once, the packets reach at MAC interface, MAC layer maps the packets to its ACs according to their priorities.

A. Source Directed Multipath Routing (SDMR)

Forward error correction (FEC) schemes are generally useful for real-time data transmission in order to avoid retransmissions since it incurs delay in each attempt. On the other hand, video streams in WMSNs have higher bandwidth requirements in contrast to the scalar data in conventional WSNs. Higher channel coding rate and larger video data size need an excessive amount of bandwidth. It is, therefore, imperative to utilize the available bandwidth and other resources efficiently to support the QoS in WMSN. To meet these objectives, we propose a multipath routing algorithm to find geographically disjoint multiple paths based on only local state information.

In multipath routing, source i determines three possible paths directed towards the destination sink s as shown in Fig. 1. P_2 is a primary path and obviously the shortest between the source and destination among the three paths. The next hop is always chosen within the distance r from the origin line, called path origin line (*PoL*) and is used as a primary path. The two secondary paths, P_1 and P_3, are also determined on location basis and restricted above and below r unit of distance from *PoL*, respectively. Hence, the three paths, P_2, P_1, and P_3, are directed along PoL, PoL' (r unit of distance above PoL) and PoL'' (r unit of distance below *PoL*), respectively. However, the separation of r is loosely restricted that may be violated when the

constraints have to be satisfied at the forwarding node. In the following, first of all, we describe the heuristic to select the next hop forwarding node and then provide an analysis for maximum hops count of these paths to determine end-to-end QoS support.

1) Forwarding Node Selection: Due to random deployment, it is highly likely that the forwarding nodes do not lie exactly on *PoL* and there is always certain displacement from *PoL*. The relaying node i, which is source initially, computes the displacement $\Delta d_{(j,PoL)}$ for the forwarding node j and uses it in the next hop selection. Let $I(x_i, y_i)$, $J(x_j, y_j)$, and $S(x_s, y_s)$ be the positions of relay node i, the potential forwarding node j, and the sink, respectively. We take $PoL = IS$ and the displacement of point J from PoL, which is the closest to PoL at the tangent to the line at an unknown point P passing through J. Therefore, the dot product of the tangent and PoL is 0, i.e., $(J-P)\cdot(S-I)=0$. Thus, we obtain the point of intersection $P(x,y)$ of the tangent on PoL as

$$x = x_i + c(x_s - x_i), \quad y = y_i + c(y_s - y_i)$$

where the value of slope c is

$$c = \frac{(x_j - x_i)(x_s - x_i) + (y_j - y_i)(y_s - y_i)}{\|S - I\|^2}.$$

Hence, the displacement of j from PoL can be obtained as

$$\Delta d_{(j,PoL)} = \sqrt{(x - x_j)^2 + (y - y_j)^2}.$$

Now, in order to select a forwarding node, node i computes the resistance (f_r) for each of the potential node that belongs to the set of forwarding nodes $\bar{N}^i$ compliance with the Constraints (6)–(8). The resistance of a node with smaller displacement is lower and higher if it is far from *PoL*. On the other hand, a node will have lower resistance if its residual energy is higher than the other forwarding candidates and vice versa. Let E_j be the residual energy of node j, $f_r(j)$ for the node j is computed as follows:

$$f_r(j) = \omega \frac{\Delta d_{(j,PoL)}}{r} - (1-\omega)\frac{E_j}{\max_{k \in \cup_{y=1}^{3} N_i^y}\{E_k\}} \tag{9}$$

where ω is a weighted factor that determines the contribution of each factor in computing the resistance force. The node $j \in \cup_{y=1}^{3} N_y^i$ is selected as a forwarding node if it has the lowest resistance, i.e., $f_r(j) = \min_{k \in \cup_{y=1}^{3} N_y^i}\{f_r(k)\}$. Thus, by combining the both factors allows to select nodes closer to *PoL* but with the balanced energy consumption. One can define a suitable value of ω according to the application requirements. If the end-to-end QoS support is far more important than the energy efficiency, then ω is kept lower that ensures that more closer nodes along *PoL* will be selected for routing and thus shorter path is followed. On the other hand, the large value of ω would enhance the network lifetime. Note that the above derivation is made for path P_2. Paths P_1 and P_3 differ by their origin lines and have longer distance. Moreover, in paths P_1 and P_3, a relay node ends up with a situation when it does not find a forwarding node with minimum advancement (a) along its path origin line. The situation is assumed to be in the neighbour of sink. Here, SDMR follows the directed routing from relaying nodes to the sink.

Note that the resistance in (9) is calculated online and does not incur any extra communication overhead since the required information are reported in the periodical *hello beacon* by each node. Thus, a forwarding node i exploits the information of its neighbors in routing by computing $f_r(j)$ for each of its neighbor $j \in N^i$ online. The computational complexity of node i is in the order of neighbor density N^i and upstream nodes N_{up}^i, i.e., $O(N^i N_{up}^i)$. Moreover, $f_r(j)$ is not computed for every packet routing decision rather periodically corresponding to the *hello beacon* frequency since the information are only updated at that frequency.

2) Maximum Hops Limit: SDMR guarantees the constraints on a single hop and provides an estimate to the maximum hops count on the basis of node density. Thus, such an assumption leads to the QoS support from source to destination subject to the certain node density. In this section, we analyze the relationship between the maximum hops count or route length (h_{is}) from source i to sink s and nodes density ρ by adopting the SDMR algorithm. By defining minimum advancement of a units of distance in the forwarding node selection, a packet always progresses a unit of distance towards the sink. Typically, the value of a is set proportional to the average distance of neighboring nodes of the source in order to incorporate node density in routing progress. In the worst scenario on path P_2 with minimum progress of a, the routing path will follow a see-saw pattern as shown in Fig. 1. The number of hops $h_{i,2}$ from source i to the sink s can be simply determined by d_{is}/a. However, it requires certain node density so that at least one forwarding node j is present at distance between a and r, i.e., $a \le (d_{is} - d_{js}) \le r$. For P_2, the forwarding region can be determined as $\pi(r^2 - a^2)/4$. Therefore, in order to meet constraint of minimum advancement a, it is required that the node density ensures the deployment of at least one node in the forwarding region. It implies that $\rho(z_2) \ge 1$, where $\rho(z_2)$ is the density in forwarding *zone 2* and is obtained as

$$\rho(z_2) = \frac{\frac{\pi}{4}(r^2 - a^2)\rho}{\pi r^2}. \tag{10}$$

For a given value of ρ, we can compute the expected route length for node i along P_2 with minimum advancement a and probability p_ρ as

$$E[h_{i,2}(a)] = p_\rho \frac{d_{is}}{a}. \tag{11}$$

Assuming uniform nodes distribution, we can compute p_ρ as

$$\begin{aligned} p_\rho &= Pr[\rho(z_2) \ge 1] \\ &= 1 - Pr[\rho(z_2) = 0] \\ &= 1 - \frac{4r^2}{\rho(r^2 - a^2)}. \end{aligned} \tag{12}$$

Hence, by putting the value of p_ρ in (11), we obtain

$$E[h_{i,2}(a)] = \left(1 - \frac{4r^2}{\rho(r^2 - a^2)}\right)\frac{d_{is}}{a}. \tag{13}$$

This gives the expected route length for P_2 at the given node density. The expected number of hops decreases as the node density increases and vice versa. Intuitively, the reliability of achieving QoS is increased if the expected route length is small

because it gives more flexible per hop delay budget (Constraint 8) and tolerable packet loss rate (Constraint 9). In order to increase the bandwidth, SDMR also explores two more node-disjoint paths P_1 and P_3. On the basis of above analysis, we can also determine the route length for paths P_1 and P_3, which is assumed to be equal, i.e., $|P_1| = |P_3|$. Unlike P_2 which is directed along the *PoL*, paths P_1 and P_3 traverse longer distance from source to destination due to the paths disjointness scheme, which is obtained $(d_{is}+2r(\sqrt{2}-1))$. Moreover, the forwarding area of P_1 and P_3 is half of the area of P_2 that requires to have node density in *zone 1* and *zone 3* twice the density in *zone 2*. Therefore, the expected hops count $(E[h_{i,13}])$ can be estimated accordingly as

$$E[h_{i,13}(a)] = Pr[\rho(z_2) \geq 2]\frac{d_{is}+2r(\sqrt{2}-1)}{a} = \left(1-\frac{8r^2}{\rho(r^2-a^2)}\right)\frac{d_{is}+2r(\sqrt{2}-1)}{a}. \quad (14)$$

Intuitively, we obtain the possible hops count $E[h_{is}]$ for SDMR by considering node density as

$$E[h_{is}] = \begin{cases} E[h_{i,13}(a)], & \text{if } \rho(z2) \geq 2 \\ E[h_{i,2}(a)], & \text{otherwise.} \end{cases}$$

Thus, per hop delay budget τ_i for source i can be computed by $\tau/E[h_{is}]$. The value of τ_i in first case is smaller due to larger route length, which is based on the higher node density. However, if the density is lower as conditioned in the second case, then the value of τ_i is higher that relaxes the per hop delay requirement due to lower node density. Hence, the routing algorithm adopts to the given node density.

Lemma 1: Paths P_1, P_2, and P_3 are node disjoint if $\rho \geq 8r^2/r^2-a^2$.

Path origin lines for the three paths have minimum separation of r units of distance. That is a node lying on one *PoL* cannot receive transmission from a node on other *PoL*. Therefore, paths can be established node disjoint if the node density is adequate to have forwarding nodes along the *PoLs*. The node density for path P_2 is derived in (11), where P_2 can be established if and only if $\rho(z_2) \geq 1$. Alternatively, we can express the condition as

$$\frac{\rho(r^2-a^2)}{4r^2} \geq 1.$$

The above inequality can be restated as

$$\rho \geq \frac{4r^2}{r^2-a^2}.$$

Since the forwarding zones z_1 and z_3 for paths P_1 and P_3, respectively, are half of the forwarding zone z_2, node density requirement for paths P_1 and P_3 is twice. Intuitively, in order for three paths P_1, P_2, and P_3, separated by distance r, to exist, the node density must be double the given in (11), hence

$$\rho \geq 2 \times \frac{4r^2}{r^2-a^2}.$$

B. Source Admission Control

In order to meet the bandwidth constraint, the framework incorporates some source admission control at each node. The admission control is based on the estimated traffic flow of the forwarding node. Recall that $R^j_{in}(t)$ and $R^j_{out}(t)$ represents the total incoming traffic from upstream nodes of j and total outgoing data rate from node j to its downstream neighboring nodes through the paths P_1, P_2, and P_3, respectively, at any time instant t. The available bandwidth of a link between nodes i and j at time t is represented by $b_{ij}(t)$. Thus

$$R^j_{out}(t) = \max_{u\in N^j_1}\{b_{ju}(t)\} + \max_{v\in N^j_2}\{b_{jv}(t)\} + \max_{w\in N^j_3}\{b_{jw}(t)\}$$

$$R^j_{in}(t+1) = R^j_{in}(t) + x_{ij}(t)r_i$$

where x_{ij} is a control function that indicates the admission of a new upstream source node i through node j. In other words, it implicitly provides a flow admission control at node j to avoid congestion and takes the value as

$$x_{ij}(t) = \begin{cases} 1, & \text{if } (r_j + R^j_{in}(t) + r_i) \leq R^j_{out}(t) \\ 0, & \text{otherwise.} \end{cases}$$

Each source i stamps the packet with its encoding data rate r_i and delay budget τ_i, and transmits it over to first hop forwarding node. Source node is generally able to find first hop forwarding node because it keeps the bandwidth reserved at outgoing links for its own requirement. In other words, node i allows the upstream traffic only if $(R^i_{in}+r_i) \leq R^i_{out}$, where its own traffic r_i may not necessarily be in progress and may start at any moment. Therefore, we can assume that the source is able to find a forwarding node for at least one hop. Source transmits only its first packet, representing the new source *request*, and waits until network feedback is received. If the *request* packet is successfully received at the sink, no matter what path is followed, then the sink generates the *ACK* packet and sends it to the source node. If a source receives an *ACK*, then it resumes the transmission of subsequent packets.

The transmission of first packet may be stopped at any intermediate relay node, which is due to the violation of any of the constraints. Mainly, it happens due to lack of required bandwidth. However, if the relay node j has adequate bandwidth to forward the packet and the mean packet delay τ_{jk} of its existing traffic to any of the forwarding node k is also smaller than τ_i, then it accepts the source request and updates its input traffic to $(R^j_{in} + r_i)$.

Although the bandwidth is reserved at the first relay node, it does not mean that it will necessarily be selected for subsequent packets. This ensures that at least one node would exist along the path to route the source traffic. In practice, any node in forwarding region can be chosen along the path vicinity within r units of distance along *PoL*. Intuitively, reservation is not meant strictly for nodes allowing the source request packet to reach the sink but it is considered as the reservation along the *PoL* path vicinity. Therefore, any node in forwarding zone along *PoL* can be the potential relay node subject to the constraints. Hence,

each node estimates its bandwidth using the bandwidth estimation technique given in [24] and broadcasts the information in its periodical *hello beacon* that also contains the *residual energy* and *location coordinates*. On the other hand, if node k is unable to meet the bandwidth requirements, then it runs the link adaptation algorithm described in the next section to accommodate the *request*. If there exists no high rate transmission mode at k to provide extra bandwidth under the given link reliability constraint, then it sends back the *reject* message to its upstream node j. Node j forwards the request to another node k' in the forwarding region and continues recursively. In this situation, there are three possibilities: first, j will find a forwarding node successfully, second, there is no such forwarding node, and third, the attempt time exceeds the hop delay τ_i. The first case lets the request to proceed further at downstream nodes. However, the second and third cases trigger the *reject* message. When the *reject* message is received by the source i, it reduces its encoding data rate according to (1) and sends the request again until either the *request* is successful or the tolerable distortion threshold has reached in reducing the encoding rate. In the former case, the source is admitted in the network and has acquired sufficient resources to transport its video. While in the latter case, the source cannot be accommodated at the moment and will initiate its request after the next bandwidth sampling interval.

C. Multirate Link Adaptation (MLA)

Link adaptation algorithm is designed to switch between different transmission modes from the set χ of size m either to increase the bandwidth or to reduce the packet error rate. In former case, the link is promoted to a higher rate mode to accommodate the new source request. While in the latter case, the link is demoted to a lower rate but more reliable transmission mode to minimize packet error rate. Assume that j is operating at mode q, for $0 \leq q < m$, that supports b_q bandwidth with link error rate χ_q. Therefore, at mode q, the expected packet error rate μ_j^q can be computed by $1-(1-\chi_q)^l$. Now let, $\tilde{\mu_j^q}$ be the packet error rate measured during the interval t_Δ seconds. The ratio $\tilde{\mu_j^q}/\mu_j^q$ gives an estimate to the deviation from the expected packet error rate, which is accounted in selection of a new transmission mode. Recall that P_L is the end-to-end tolerable packet loss rate that accounts for the total packet error rate over h_{is} number of hops from source i to the sink s. In order to keep the packet error rate lesser than the given threshold value P_L, the link selection must ensure that $(\mu_j^q)^{h_{is}} \leq P_L$ since the packet loss rate is multiplicative at each hop. In other words, the link quality at each forwarding node j for traffic of source i ensures that $\mu_j^q \leq \sqrt[h_{is}]{P_L}$. Besides the current request, the reliability of existing sources must also be maintained. Hence, we can define the minimum reliability threshold of node j for set of upstream sources $\aleph_j$ as

$$\mu_{th}^j = \min_{k \in \aleph_j} \left\{ \sqrt[h_{ks}]{P_L},\ \sqrt[h_{is}]{P_L} \right\}. \tag{15}$$

Let t_ϵ^j and t_{oh} be the medium access time of node j and the overhead time during a sampling period of length t_Δ, respectively. The available bandwidth is then proportional to $t_\epsilon^j - t_{oh}/t_\Delta b_q$. If it is smaller than the required data rate r_i of new request, then the request can be accepted if there exists a mode q' subject to the following conditions:

1) $q < q' < m$
2) $\max\{\tilde{\mu_i^q}/\mu_i^q \mu_i^{q'}, \mu_i^{q'}\} \leq \mu_{th}^j$
3) $r_i \geq (b_{q'} - b_q)(t_\epsilon^j - t_{oh}/t_\Delta)$

The first condition implies that there exists a higher rate transmission mode. The second condition ensures that per hop reliability is not compromised to accommodate the new source. While in the third condition, we switch to a mode that can provide enough bandwidth to accept the request. Intuitively, if none of the condition is unsatisfied, the link remains operating on its current mode and the *request* is rejected. It is important to note that the selection of a new mode must not degrade the reliability of the existing sources in terms of error rate.

D. IEEE 802.11e Based MAC

In this section, we describe how per hop delay is guaranteed at MAC layer, which is stamped in the packet as hop delay budget by the routing protocol SDMR. As discussed earlier, MAC layer is based on the enhanced distributed coordination function (EDCF) provided in IEEE 802.11e wireless LAN standard to support QoS demanding traffic [31]. Recall that QoS is supported in this standard through service differentiation by providing four prioritized access categories AC_VO, AC_VI, AC_BE, and AC_BK for voice, video, best-effort, and background traffic, respectively. This implies that the frames queued in AC_VO get higher priority than frames in the other ACs. We exploit this prioritized scheduling to support QoS in more efficient way. Furthermore, the standard also defines the size of burst called $TXOP_{limit}$ that allows to transmit $TXOP_{limit}$ number of frames simultaneously through single RTS/CTS frames exchange.

Besides the prioritized scheduling, we also define a maximum burst size γ to set different $TXOP_{limit}$ for each AC in which AC_VO is allowed the largest burst size and AC_BK has the smallest value according to their priority. Hence, the value of $TXOP_{limit}[AC]$ for access category AC is set to γ/α, where α takes the values 1, 1, 4, and 8 for AC_VO, AC_VI, AC_BE, and AC_BK, respectively. The value of γ should be greater or equal to the highest value of α so that at least one frame can be transmitted from the lowest priority AC in each round. The values of α are not fixed and can be set according to application requirement. However, such a scheme would prevent starvation of the lower priority traffic and intuitively provides fairness corresponding to the values of α. Let p_c be the collision probability [25] and μ_q be the packet error rate at mode q. Assuming that frame failures due to collision and transmission error are independent, the probability of failure p_f is $p_c + \mu_q$. The number of expected retransmissions $\phi(q)$ at link operating in mode q can be obtained by $\phi(q) = \left\lceil \frac{1}{1-p_f} \right\rceil$.

The probability p_s of successful frame transmission after $\phi(q)$ attempts is $(p_f)^{\phi(q)}(1-p_f)$ before reaching maximum retransmission limit Q, i.e., for $\phi(q) < Q$. However, it is zero

for $\phi(q) > Q$ and $(p_f)^Q$ when $\phi(q) \cong Q$. We can obtain the average backoff time for an access category c as follows:

$$t_{bo}(c) = \sum_{k=0}^{\phi} p_s \frac{\min(2^k CW_{\min}[c], CW_{\max}[c]) - 1}{2} t_{slot}.$$

Hence, the time t_x required for transmission of a frame is

$$t_x(c) = t_{bo} + AIFC[c] + t_{rts} + t_{cts} + t_{data} + t_{ack} + 3 \times SIFS.$$

However, this time is, generally, required for a frame at the head of the queue of access category c contending for the medium. In our approach, each frame does not spend this time to transmit the packet because a burst of frames of size $TXOP_{limit}[c]$ are transmitted consecutively once the medium is captured for the first frame in burst. Therefore, the second frame in burst has to wait $t_{data} + t_x$ sec, the second frame $2t_{data} + t_x$ sec, and so on. The burst transmission time t_α is important to determine the packet delay, which is obtained as

$$t_\alpha(c) = (\alpha[c] - 1)t_{data} + t_x[c].$$

Frames in the next burst from the same AC will have to wait for the next round until each AC gets its fair share of the medium. Hence, the worst case expected delay can be computed as the accumulative delay required for transmission of a burst in each AC having frames for transmission:

$$E[t_{hop}] = \sum_{c=0}^{3} Pr[|c| > 0] t_\alpha[c] \tag{16}$$

where $Pr[|c| > 0]$ represents the probability of cth AC having frames for transmission. Thus per hop delay requirement for source i is met when $\tau_i \leq E[t_{hop}]$ at the forwarding node j. If we assume only multimedia traffic, then frames are only present in AC_{VO} and AC_{VI} and we take $Pr[|AC\_BE| > 0] = Pr[|AC\_BK| > 0] = 0$. Thus the expected delay is significantly reduced. The source admission control algorithm at each node accepts the source if it can tolerate the delay $E[t_{hop}]$. In our proposed enhancements in IEEE 802.11e, **I**-Frames, *reject, feedback*, and *Ack* messages are assigned to AC_VO, while only **W**-Frames are initially assigned to AC_VI. However, if there exists a **W**-Frame that cannot be transmitted in the current burst but has per hop delay limit shorter than $E[t_{hop}]$, then it is moved to AC_VO. Thus, it allows to transmit short delay constrained **W**-Frames in the following burst since the control is switched to AC_VO and intuitively it gets opportunity on both ACs.

To further minimize the chances of missing per hop deadline, the scheduling algorithm selects a frame that has minimum per hop remaining deadline computed locally using hop delay budget rather than a head-of-queue frame and passes it from link layer queue to MAC for transmission. This ensures that nodes which are at higher hops far from the sink and has, therefore, shorter deadlines, get higher priority within the same AC and are transmitted earlier. Such a scheme would balance the packet delivery of farther and closer nodes and implicitly provides bandwidth fairness. The average frame delay in each AC at different values of γ is shown in Fig. 3. The larger the burst size is allowed, lower the average delay is experienced. However, a new frame, which cannot be transmitted in the current burst of any AC, experiences delay as sum of the delay in all ACs or time taken to complete a round. Thus, as we increase the size of bursts by choosing larger value of γ, the more time will be required to complete the round. Hence, the maximum frame transmission delay increases accordingly. The analysis provide a basis to select the suitable parameters according to the given application scenario and requirement. For example, taking the farthest possible source i from the sink, we can determine the worst per hop delay budget τ_i at given node density. Using this value of τ_i, we can choose the value of γ for which $\tau_i \leq E[t_{hop}]$. Intuitively, the per hop delay is guaranteed in our IEEE 802.11e based MAC protocol.

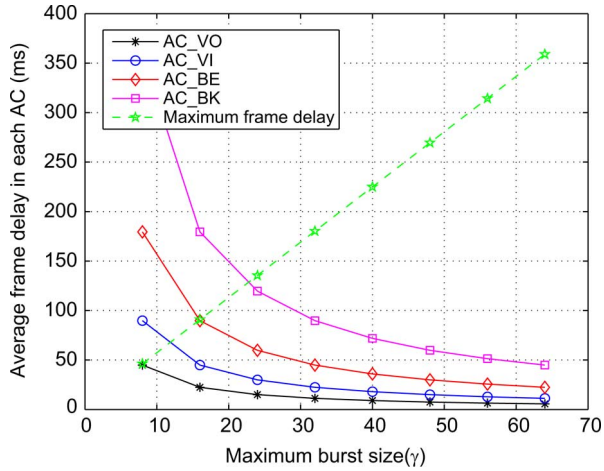

Fig. 3. Average frame delay in each AC with different values of γ and fixed collision probability 0.2.

Lemma 2: End-to-end QoS can be guaranteed with some probability p, given that $\rho \geq 4r^2/(1-p)(r^2 - a^2)$.

We have defined the video distortion, bandwidth, packet delay and packet losses metrics as the QoS parameters for multimedia support in WSNs. In the proposed cross-layer architecture, source admission control algorithm ensures that per hop input traffic does not exceed the capacity of the link and thereby bandwidth is reserved for the source streams that guarantees bandwidth constraint. MLA provides link adaptation to accommodate new sources if the existing link bandwidth is inadequate. However, the algorithm switches to higher link rate if the expected packet loss rate at the new mode does not exceed the per hop packet loss tolerance of the sources and hence limits the loss rate to achieve the required packet reliability of each source. The QoS-based MAC protocol provides an estimate of the probable per hop frame delay $E[t_{hop}]$, which is used by source admission to make sure that the per hop delay requirement of source is also supported by the node. Thus, the delay constraint can be meet subject to the following condition:

$$\frac{\tau}{|P_1|} \geq t_{hop} \tag{17}$$

where the route length for P_1 can be determined based on the node density, advancement a and the maximum distance between the farthest node and sink F by $|P_1| =$

$F/aPr[\rho\pi(r^2-a^2)/4 \geq 1]$. Therefore, the inequality (17) yields

$$\frac{F}{a}Pr\left[\rho\frac{\pi(r^2-a^2)}{4}\geq 1\right]\geq\frac{\tau}{t_{hop}}$$
$$\text{or}\quad\left(1-\frac{4r^2}{\rho(r^2-a^2)}\right)\geq\frac{\tau a}{t_{hop}F}. \tag{18}$$

Here τ, a, and F are fixed values, while t_{hop} varies dynamically whose expected value $E[t_{hop}]$ is formulated in 17. The fraction on the right-hand side of the inequality (18) approximates the probability that the delay constraint is preserved, i.e., $p = Pr[t_{hop}F \leq \tau a]$. For instance, if the application sets the deadline τ too short against $E[t_{hop}]$, then it conceives low probability of achieving QoS accordingly and vice versa. Thus, (18) yields

$$\left(1-\frac{4r^2}{\rho(r^2-a^2)}\right)\geq p. \tag{19}$$

This is simplified as

$$\rho\geq\frac{4r^2}{(1-p)(r^2-a^2)}. \tag{20}$$

Hence, end-to-end QoS can be supported with the probability p subject to the constraint $\rho \geq 4r^2/(1-p)(r^2-a^2)$.

E. Sink-Feedback Based Distortion Control

Distortion is the basic metric to determine the quality of video. It appears in the video due to various factors, such as encoding losses at the source, inadequate network support for encoded data transmission and decoding losses at the receiver. Among these, network support, particularly in WSNs, is highly dynamic that causes the distortion to vary over space and time scales, which must need to be controlled to support persistent video quality. The sink-feedback based distortion control algorithm aims to keep the source video distortion under the application specified threshold Θ. The distortion threshold is used to determine the tolerable packet losses and GOP size [23], [26]. While the packets deadline is specified to account the playout delay at the sink. Therefore, if the packets are received later than their given deadline, then they are not played out and have the effect on video quality similar to lost packets. Hence, for a given value of Θ, the other parameters GOP size (g), packet loss rate (P_L), and deadline (τ) are determined. This allows to dynamically change the values of g, P_L, and τ according to varying network conditions so that the perceived video distortion is kept under Θ.

Assume that source i is sending its video to the sink and set the values of g using (1), P_L using (2) and τ corresponding to the play out delay to satisfy Θ. The sink perceives distortion D_i for the source i and the video quality is satisfactory given that $D_i \leq \Theta$. However, if $D_i > \Theta$, then the causes of higher distortion is determined by the sink. One of the possible reasons is the late packets arrival, i.e., packets are received later than τ seconds, which are treated as lost packet by the decoder and incur distortion. If the late arrivals persists for three consecutive sampling intervals, then the playout time is increased corresponding to the excessive delay of packets by the sink to remedy the impact of delayed packets. An increase in play out time would required larger buffer capacity at the sink, which is not a big deal for the sink since it does not face resource constraints problem.

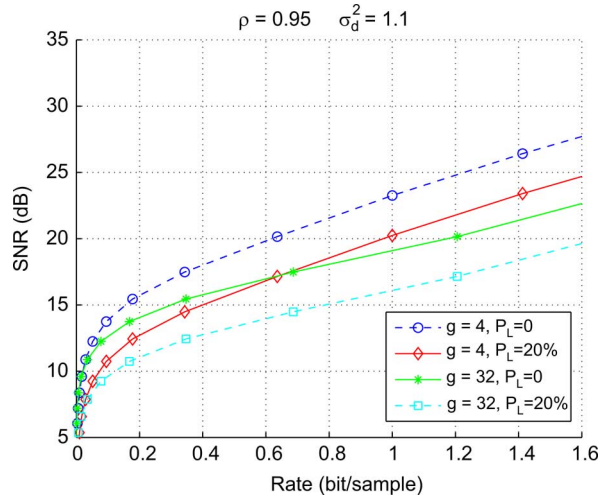

Fig. 4. Video quality measured at different GoP size g and packet loss rate P_L for 10 sources.

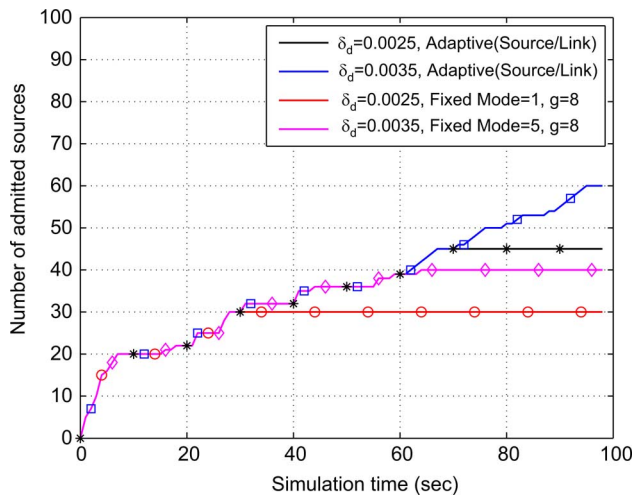

Fig. 5. Number of sources admitted at different time under the constrained distortion of each individual source.

Similarly, if the distortion is observed due to higher packet loss rate than the given value of P_L, then packet losses can be either transient or persistent. Transient losses generally occur due to temporary channel fading or noise at some intermediate nodes, which may result in higher losses temporarily. While persistent losses are observed due to network saturation that results in congestion at some intermediate nodes. Sink node observes the losses and if the loss rate does not exceed the threshold for more than three consecutive sampling intervals, then the losses are assumed transient. In the latter case, the higher loss rate is observed for more than three successive intervals and is assumed to be persistent, which is considered as a congestion indication. In either case, the sink sends a feedback message back to the source(s) for which higher losses are observed. Note that, the frequency of such messages is usually low and are transmitted with the highest priority traffic class in IEEE 802.11e protocol.

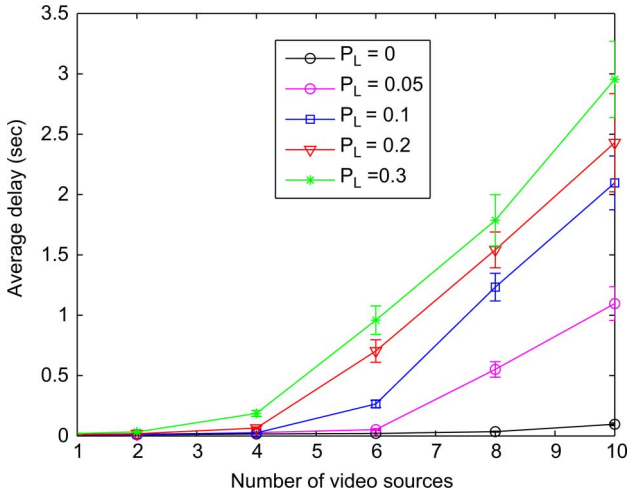

Fig. 6. Average packet delay for different number of source nodes by varying packet error rate χ.

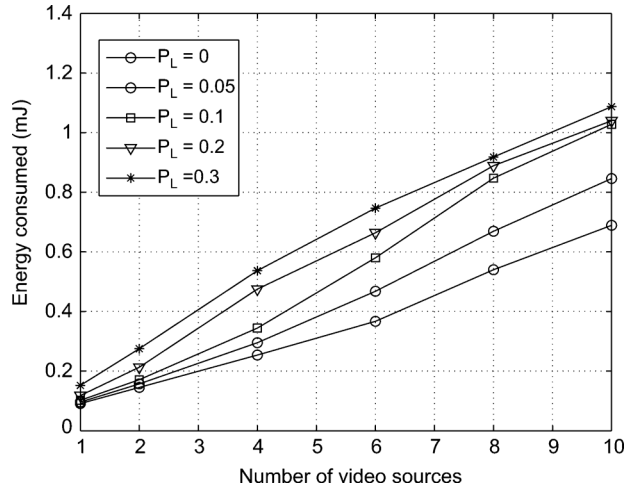

Fig. 8. Total energy consumed in sending *foreman* video stream by varying number of source nodes at different packet error rate χ.

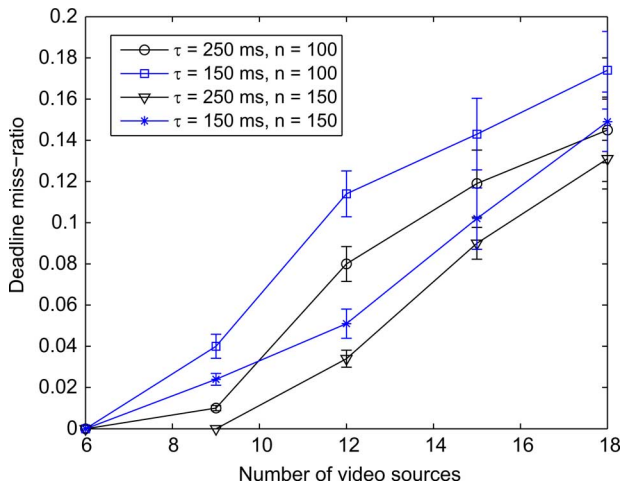

Fig. 7. Packet deadline-miss-ratio at different values of τ and node density ρ using number of nodes n at $\chi = 5\%$ by varying sources.

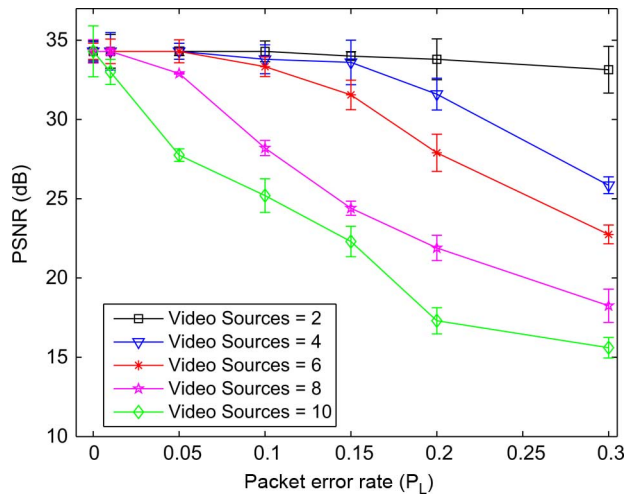

Fig. 9. Average PSNR measured at sink node by varying packet error rate χ and different number of source nodes.

The feedback message contains the observed distortion and the type of loss, which is either transient or persistent. In case of transient losses, the transmission conditions are assumed to be hostile that requires more reliable transmission mode. Thus, the source node reduces its packet loss threshold accordingly in the subsequent packets. This will trigger the link adaptation algorithm at the intermediate nodes to select a more reliable transmission mode meeting the source reliability requirements. If the persistent loss is reported by the sink, then the network capacity is presumed to be inadequate and therefore source reduces its encoding rate. This is achieved by increasing the value of g using (1) to adjust to its minimum possible rate subject to the distortion threshold. Intuitively, the distortion is controlled by taking feedback message from the sink.

V. Performance Evaluation

In this section, we evaluate the performance of the proposed cross-layer design of the routing protocol in WMSNs using network simulator *ns-2* [37] and a patch for video streaming [35]. The performance metric is the quality of video measured as the peak signal-to-noise ratio (PSNR) and mean opinion score (MOS) as a video quality of experience widely used in literature [4], packet delay, energy consumption, and throughput. The setup consists of 100 nodes deployed uniformly in 300×300 field with sink at the center. The number of sources are variable and selected at least 3 hops away from the sink. Two types of data sources are running; CBR source generates traffic at 125 *KB/sec* and video generates frames at 30 frames per second. Path loss exponent is set to 2 and shadowing deviation to 4. Simulations are performed about 5 to 10 times for each scenario with the confidence interval of 99%.

In the first scenario, the quality of video is measured for 10 video sources sending their video data to the sink. Fig. 4 shows the PSNR value at different encoding rate and GOP size g. The value of g is optimal for the distortion threshold and encoding bit rate as given in [26]. However the video quality drops if packet losses are incorporated in measuring the distortion. Hence, for higher packet losses, we reduce the value of g to achieve the desired quality. Approximately 18% PSNR value is dropped at 20% packet losses. Similarly, the adaptivity of the cross-layer design is depicted in Fig. 5 for varying number of sources up to 75. It shows that the number of admitted sources are increased

Fig. 10. Snapshot of a *foreman* video transmitted by a sensor node to sink with non-adaptive and adaptive cross-layer framework. (a) Original image in a snapshot of *foreman* video. (b) Image of a snapshot of *foreman* video with adaptive cross-layer design. (c) Image of a snapshot of *foreman* video with non-adaptive transmission.

to 40% with adaptive source and link as compared to the fixed source encoding and single mode transmission. The value can be increased further if the decoding distortion threshold δ_d is set to higher value, which is increased from 0.0025 to 0.0035. Thus by compromising the video quality to some extent, the number of admitted sources can be maximized.

In the second scenario, different number of sources are sending their *foreman* video streams at different packet error rates. The average delay of up to 10 sources is reported in Fig. 6. If the packet error rate is negligible, then there is a slight increase in delay because the link can be promoted to higher mode to attain more bandwidth and also SDMR follows multiple disjoint paths to distribute the load among three different paths. Hence, the admission of a new source does not necessarily increase the delay of existing admitted sources. On the other hand, by increasing the error rate, the delay also increases for higher number of sources. It is due to the fact that at higher error rate, the lower reliable mode of lesser bandwidth is selected that causes to increase the queuing delay and thus the packet transmission delay is increased. The delay for six video streams at 30% packet error rate is below 1 second. Thus, setting playout deadline to 1 second would limit the distortion due to higher packet error rate.

Similarly, Fig. 8 illustrates the total energy consumed in sending video streams by different number of sources. Obviously, energy increases by increasing the traffic that is due to the admission of new sources as well as retransmission due to higher packet error rate. However the energy consumption can be controlled for retransmissions using link adaptation. For higher error rate with smaller number of sources, the link is switched to more reliable mode to conceal the higher error rate and thus lesser energy is consumed, which is evident up to four sources. Thus by increasing the error to 30% for 10 sources, approximately 40% more energy is consumed. Moreover, the load balancing strategy of SDMR prevents from continuously selecting the same forwarding node and hence the deviation in the remaining energy of nodes is observed to be very small. Fig. 7 depicts the packet deadline miss-ratio. At lower node density at $n = 100$, miss-ratio is high particularly at lower deadline $\tau = 150$ ms. However, the miss ratio is reduced by setting the required probability to higher value of 0.85. Following Lemma 2, the density is increased by increasing the nodes from 100 to 150 to decrease the miss-ratio. Moreover, when τ is higher, per hop delay budget is also higher at the given node density and intuitively miss-ratio is lower.

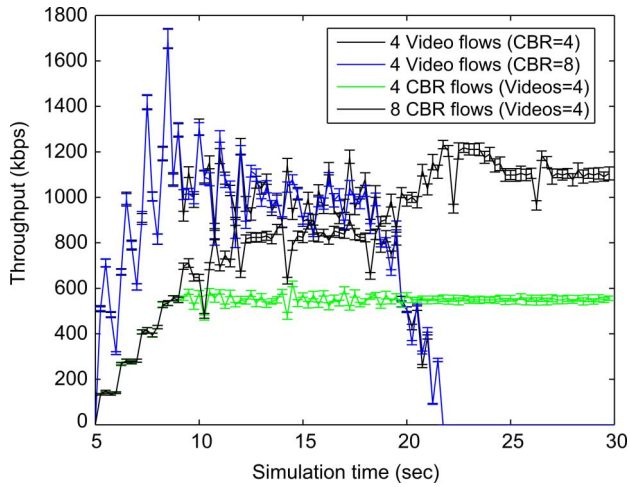

Fig. 11. Aggregated throughput of video and CBR flows.

The effect of packet error rate is also investigated in measuring the quality of video in terms of PSNR, where the threshold is set to 15 dB. The sources are admitted only if the sink perceives PSNR value greater than the defined threshold. Fig. 9 shows the PSNR value for different sources at varying packet error rate. Whenever a source is admitted, it causes to increase in delay and possibly selection of lower reliable mode that eventually degrades the quality. The admission is made as long as the sink does not complain for the lower PSNR value. As a result, higher error rate keeps degrading the quality until the threshold is reached. Increasing packet error rate to 30%, the PSNR is dropped to approximately 50% for 10 sources but is not much affected at lesser number of sources. The snapshot of the video is shown in Fig. 10. It can be seen that the cross-layer framework preserves the quality of video in Fig. 10(b) to some extent by differentiating **I** and **W** frames, which is closer to the original snapshot in Fig. 10(a), while the quality in non-adaptive design is faded as shown in Fig. 10(c).

Throughput of the framework is also reported in Fig. 11. The impact of CBR traffic on video flows is studied to observe the throughput of video traffic. It can be seen that the throughput of video traffic is not much affected by increasing the CBR flows from 4 to 8 that remains persistent. The high peaks in video flows are mainly caused due to the arrival of **I**-frames of

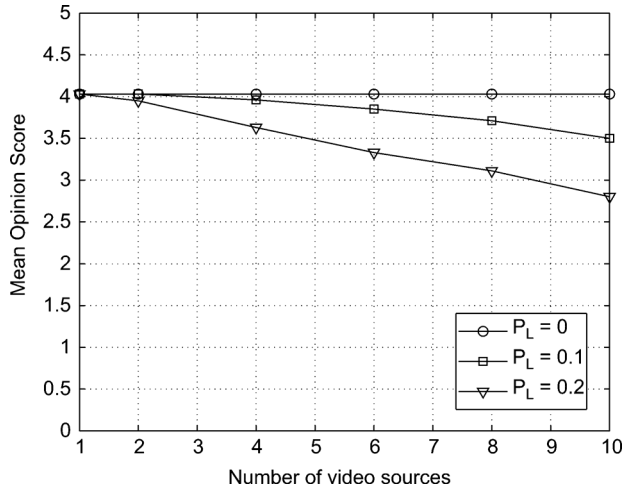

Fig. 12. MOS values at different packet error rate by varying the number of video sources.

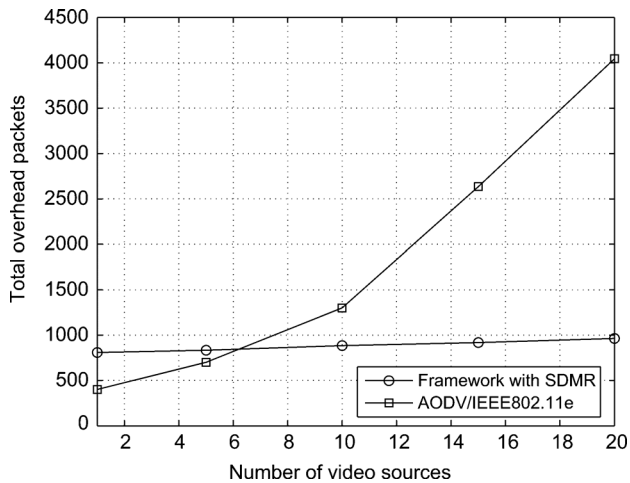

Fig. 13. Overheads in the proposed framework for different number of sources.

larger size. However as the video streams are finished at 20 seconds of time, the throughput of CBR flows increases suddenly to achieve its maximum required rate.

MOS is a subjective metric to measure quality-of-experience (QoE) of a video at the application level. This metric of the human quality impression is usually given on a scale that ranges from 1 (worst) to 5 (best). In this framework, the PSNR of every single frame is approximated to the MOS scale using the mapping function as Excellent (PSNR $>$ 37 dB) $=$ 5, Good (37 $\geq$ PSNR $\geq$ 31 dB)$=$ 4, Fair (31 $>$ PSNR $\geq$ 25 dB)$=$ 3, Poor (25 $>$ PSNR $\geq$ 20 dB)$=$ 2, and Bad (PSNR $<$ 20 dB)$=$ 1. Fig. 12 shows MOS for the *foreman* video that reveals the degradation of video quality from Good to Fair by increasing the error rate to 20% for sources above 8, while it sustains the same quality for error rate lower than 20% and smaller sources regardless of error rate. The lower error rate can be concealed by the framework using adaptive link and routing techniques but higher packet losses at larger video sources exceed the possible link capacity.

Finally, the overheads incurred during data delivery in the proposed framework is shown in Fig. 13. Clearly, the number of overhead packets does not increase with the same proportion of increase in the number of sources in our framework employing SDMR routing, which is barely 10% by increasing sources from 1 to 20. On the other hand, the increase in overhead packets is significantly high when AODV is used with IEEE 802.11e MAC protocol. It is due to the fact that SDMR transmits periodic *hello* messages regardless of the number of sources showing its control messages higher than AODV initially but maintains with slight increase afterwards due to some reporting messages.

VI. Conclusions

This paper proposed a cross-layer design architecture to address the problem of maximizing the number of video sources admitted without affecting the quality of existing sources in WMSNs. In this architecture, application layer, routing layer and data link layer interact adaptively with each other to achieve the objective. Unlike the existing studies, our cross-layer design is capable of interacting with the application layer to choose the appropriate GoP size according to network conditions and feedback received from the sink. Moreover, we propose a novel source directed routing protocol, which is a stateless routing and does not negotiate the end-to-end resource requirement. QoS is provided hop by hop, which is achieved end-to-end by bounding the maximum hop count on the basis of the node density. The design also exploits the availability of multiple transmission modes of wireless interface, i.e., IEEE 802.11, to select a mode that has higher transmission rate with constrained error rate. Hence, the value of g, routing paths and transmission modes are selected such that the distortion in video is restricted below the application defined threshold with maximum number of sources sending their data to the sink.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Net*, vol. 51, no. 4, pp. 921–960, Mar. 2007.

[2] *Amendment 8 Medium Access Control (MAC) Quality of Service (QoS) Enhancements*, IEEE Std. 802.11e, Jul. 2005.

[3] T. Melodia and I. F. Akyildiz, "Cross-layer QoS-Aware communication for ultra wide band wireless multimedia sensor networks," *IEEE J. Select. Areas Commun.*, vol. 28, no. 5, pp. 653–663, 2010.

[4] M. A. Igartua, L. J. C. Llopis, V. C. Frias, and E. S. Gargallo, "A game-theoretic multipath routing for video-streaming services over mobile Ad Hoc networks," *Comput. Net*, vol. 55, no. 13, pp. 2985–3000, 2011.

[5] L. Shu, M. Hauswirth, L. Wang, Y. Zhang, and J. H. Park, "Cross-layer optimized data gathering in wireless multimedia sensor networks," in *Proc. CSE'09*, Oct. 2009, pp. 961–966.

[6] N. Saxena, A. Roy, and J. Shin, "Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks," *Comput. Net*, vol. 52, no. 13, pp. 2532–2542, 2008.

[7] O. Farrag, M. Younis, and W. D. Amico, "MAC support for wireless multimedia sensor networks," in *Proc. IEEE GLOBECOM'09*, 2009, pp. 3781–3786.

[8] H. Kim and S.-G. Min, "Priority-based QoS MAC protocol for wireless sensor networks," in *Proc. IEEE IPDPS'09*, 2009, pp. 1–8.

[9] M. Chen, V. C. Leung, S. Mao, and Y. Yuan, "Directional geographical routing for real-time video communications in wireless sensor networks," *Comput. Commun.*, vol. 30, no. 17, pp. 3368–3383, 2007.

[10] J. B. Othman and B. Yahya, "Energy efficient and QoS based routing protocol for wireless sensor networks," *J. Parallel Distrib. Comput.*, vol. 70, no. 8, pp. 849–857, 2010.
[11] E. Felemban, S. Member, C. Lee, and E. Ekici, "MMSPEED: Multipath multi-SPEED protocol for QoS guarantee of reliability and timeliness in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 6, pp. 738–754, 2006.
[12] S. Li, R. Neelisetti, C. Liu, S. Kulkarni, and A. Lim, "An interference-aware routing algorithm for multimedia streaming over wireless sensor networks," *Int. J. Multimedia Appl.*, vol. 2, no. 1, pp. 10–30, 2010.
[13] X. Huang and Y. Fang, "Multiconstrained QoS multipath routing in wireless sensor networks," *Wireless Net*, vol. 14, no. 4, pp. 465–478, 2008.
[14] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Comput. Net*, vol. 45, no. 3, pp. 245–259, 2004.
[15] F. Oldewurtel, J. Ansari, and P. Mahonen, "Cross-layer design for distributed source coding in wireless sensor networks," in *Proc. SENSORCOM'08*, 2008, pp. 435–443.
[16] P. F. James, C. How, D. Bull, and A. Nix, "Distortion-based link adaptation for wireless video transmission," *EURASIP J. Adv. Signal Process.*, vol. 2008, p. 17, 2008.
[17] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, pp. 80–94, 2004.
[18] M. V. Schaar and D. S. Turaga, "Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission," *IEEE Trans. Multimedia*, vol. 9, no. 1, pp. 185–197, 2007.
[19] I. Haratcherev, J. Taal, K. Langendoen, R. Lagendijk, and H. Sips, "Optimized video streaming over 802.11 by cross-layer signaling," *IEEE Commun. Mag.*, vol. 44, no. 1, pp. 115–121, 2006.
[20] Z. He, W. Cheng, and X. Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization," *IEEE Trans. Circuits Syst.*, vol. 18, no. 5, pp. 596–608, 2008.
[21] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraints," *IEEE Trans. Circuits Syst.*, vol. 15, no. 5, pp. 645–658, 2005.
[22] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *EURASIP Signal Process.*, vol. 86, no. 11, pp. 3123–3130, Nov. 2006.
[23] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.
[24] C. Sar, C. Chaudet, G. Chelius, and I. G. Lassous, "Bandwidth estimation for IEEE 802.11 based Ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 10, pp. 1228–1241, Oct. 2008.
[25] C. T. Chou, K. G. Shin, and S. N. Shankar, "Contention-based airtime usage control in multirate IEEE 802.11 wireless LANs," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 1179–1192, 2006.
[26] M. Tagliasacchi, L. Frigerio, and S. Tubaro, "Rate-distortion analysis of motion-compensated interpolation at the decoder in distributed video coding," *IEEE Signal Process. Lett.*, vol. 14, no. 9, pp. 625–628, 2007.
[27] X. San, H. Cai, J. G. Lou, and J. Li, "Multiview image coding based on geometric prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1536–1548, 2007.
[28] L. Zhou, X. Wang, W. Tu, G. Mutean, and B. Geller, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE J. Select. Areas Commun.*, vol. 28, no. 3, pp. 409–419, 2010.
[29] H.-P. Shiang and M. Var der Schaar, "Distributed resource management in multi-hop cognitive radio networks for delay sensitive transmission," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 941–953, 2009.
[30] P. Frossard, J. C. De Martin, and R. Civanlar, "Media streaming with network diversity," *Proc. IEEE*, vol. 96, no. 1, pp. 39–53, 2008.
[31] Y. P. Fallah, P. Nasiopoulos, and H. Alnuweiri, "Efficient transmission of H.264 video over multirate IEEE 802.11e WLANs," *EURASIP J. Wireless Commun. Net*, vol. 2008, p. 14, 2008.
[32] D. Guinard, V. Trifa, T. Pham, and O. Liechti, "Towards Physical Mashups in the Web of Things," in *Proc. INSS'09*, 2009, pp. 1–4.
[33] Web of Things, Dec. 27, 2011. [Online]. Available: http://www.webofthings.org/.
[34] M. Geoff, "The 6LoWPAN architecture," in *Proc. ACM EmNets'07*, 2007, pp. 78–82.
[35] C. Ke, C. Shieh, W. Hwang, and A. Ziviani, "An evaluation framework for more realistic simulations of MPEG video transmission," *J. Inf. Sci. Eng.*, vol. 440, no. 2, pp. 425–440, 2008.
[36] G. Mao, B. Anderson, and B. Fidan, "Path loss exponent estimation for wireless sensor network localization," *Comput. Netw.*, vol. 51, pp. 2467–2483, 2007.
[37] UC Berkeley, LBL, USC/ISI and Xerox PARC, The Network Simulator ns-2, The VINT Project. [Online]. Available: http://www.isi.edu/nsnam/ns/.

Ghalib A. Shah (M'09) received the Ph.D. degree in computer engineering from Middle East Technical University, Çankaya, Turkey, in 2007 and the B.Sc. degree with major math in 1997 and the M.Sc. degree in computer science in 2000, both from Bahauddin Zakariya University, Multan, Pakistan.

In 2007, he joined the National University of Science & Technology, College of E & ME, as an Assistant Professor. Later, he joined the Center for Advanced Research in Engineering (CARE), Islamabad, as a member Technical Staff and lead various networking projects. He was also awarded a COMSTECH-TWAS joint research grant for young researchers. Recently, he is working at the Next Generation Wireless Communications Laboratory, Koc University, Istanbul, Turkey, as a PostDoc fellow. His research interests include the design and analysis of communication protocols from MAC to transport layer for diverse domain of wireless networks.

Weifa Liang (M'99–SM'01) received the B.Sc. degree from Wuhan University, China, in 1984, the M.E. degree from the University of Science and Technology of China in 1989, and the Ph.D. degree from the Australian National University in 1998, all in computer science.

He is currently an Associate Professor in the Research School of Computer Science at the Australian National University, Canberra, Australia. His research interests include design and analysis of energy-efficient routing protocols for wireless ad hoc and sensor networks, information processing in wireless sensor networks, routing protocol design for WDM optical networks, design and analysis of parallel and distributed algorithms, combinatorial optimization, and graph theory.

Ozgur B. Akan (M'00–SM'07) received the Ph.D. degree in electrical and computer engineering from the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, in 2004.

He is currently a full professor with the Department of Electrical and Electronics Engineering, Koc University, Istanbul, Turkey, and the director of the Next-Generation and Wireless Communications Laboratory. His current research interests are in wireless communications, nano-scale and molecular communications, and information theory.

Dr. Akan is an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *International Journal of Communication Systems* (Wiley), and *Nano Communication Networks Journal* (Elsevier). He is currently General Co-Chair of ACM MOBICOM 2012, IEEE MoNaCom 2012, and TPC Co-Chair of IEEE ISCC 2012.