A Survey on Energy-Efficient Routing Techniques with QoS Assurances for Wireless Multimedia Sensor Networks

Samina Ehsan and Bechir Hamdaoui

School of EECS
Oregon State University
{ehsan,hamdaoui}@eecs.oregonstate.edu

***Abstract*— The recent technological advances in micro electro-mechanical systems have promoted the development of a powerful class of sensor-based distributed intelligent systems capable of ubiquitously retrieving multimedia information, namely Wireless Multimedia Sensor Networks (WMSNs). WMSNs are gaining more popularity day by day as they are envisioned to support a large number of both non-real time and real-time multimedia applications. However, satisfying the stringent quality of service (QoS) requirements of multimedia transmission in a resource-constrained sensor network environment places new challenges to routing. As an outcome, optimal energy and application-specific QoS aware routing for WMSNs has gained considerable research attention recently. In this paper, current state-of-the-art in energy-efficient routing techniques for WMSNs is surveyed together with the highlights of the performance issues of each strategy. We outline the design challenges of routing protocols for WMSNs followed by the limitations of current techniques designed for non-multimedia data transmission. Further, a classification of recent routing protocols for WMSNs and a discussion of possible future research trends are presented.**

I. Introduction

Advances in embedded microprocessors, low-power analog and digital electronics, and radio communications have enabled the development of small and low-priced sensor nodes (SNs) that made wireless sensor networks (WSNs) one of the promising technologies during the past decade. In most cases, a WSN is comprised of a large number of irreplaceable, battery-powered SNs, scattered densely and randomly in a geographical area of interest. In general, the SNs in a WSN sense and gather data from surrounding environment and transmit it to logically more potent nodes, called sinks, to perform more intricate processing. Sensor based applications span a wide range of areas, including scientific research, military, disaster relief and rescue, health care, industrial, environmental, and household monitoring. Recently, researchers have realized that to extract more realistic and precise information of the fast changing events in the real world, the abilities of traditional sensor nodes should be enhanced. So evidently, the growing pace of technological demand has worked as the driving force for designing sensors capable of sensing and producing multimedia data. The availability of low cost and also miniature size cameras or microphones made it possible that sensor nodes fitted with them can extract more descriptive information about the ambience. These nodes form more powerful, distributed systems, conferred as wireless multimedia sensor networks (WMSNs). That is, networks of wirelessly connected smart devices capable of capturing video and audio streams, still images, and scalar sensor data pervasively both in real-time and non-real-time. As shown in Fig. 1, based on the requirements of different applications, the architecture of WMSNs can be divided into three different classes: single layered, flat and homogeneous; single layered, clustered and heterogeneous; and multi-layered, clustered and heterogeneous. In single layered, flat and homogeneous architecture, every multimedia sensor node has the same physical capabilities, and a subset of the deployed sensors has higher processing capabilities termed as multimedia processing hubs, which are responsible for in-network processing (i.e., data aggregation, discard of redundant data) in a distributed fashion and relaying to a sink through a multi-hop path [1]. Single layered, clustered and heterogeneous architecture is composed of heterogeneous sensory nodes (multimedia sensor nodes, basic WSN or scalar sensor nodes, etc.), where the sensor nodes in the cluster gather sensory information from the environment and then send it to the cluster head, which has more resources and computational power as compared to other cluster nodes. The processed information is then transmitted to the sink via wireless link by the cluster head. The multi-layered, clustered and heterogeneous architecture is comprised with a layer of scalar sensor nodes for performing simple tasks of gathering the scalar information from the surrounding ambience, a layer of medium resolution multimedia sensor nodes capable of gathering multimedia information from the surrounding environment, and a final layer of highly powerful multimedia sensor nodes for performing complex tasks and transmitting data to the sink.

Most deployed WSNs measure physical phenomena like temperature, pressure, humidity, or location of objects. In general, most of those applications have low

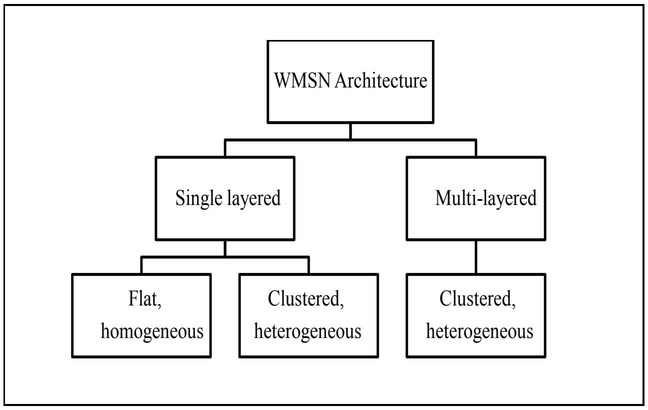

Fig. 1: Classification of the architecture of WMSNs

bandwidth demands, and are usually delay tolerant [1]. On the other hand, WMSNs, comprised of sensor devices equipped with audio and visual information collection modules, can have the ability to retrieve multimedia data, store or process data in real-time, correlate and fuse multimedia data originated from heterogeneous sources, and wirelessly transmit collected data to desired destinations [1]. Moreover, WMSNs are designed for those real-time applications which demand strict deadline, low delay, high throughput, and reliability as well as those non-real time applications which require high or medium bandwidth, loss intolerance etc. However, achieving energy efficiency is the common challenge for both networks.

WMSNs have not only stretched the horizon of traditional sensor networks, but also proliferated significantly to a variety of novel applications. Recent years have witnessed the pilot deployments of WMSNs for a class of real-time mission-critical and monitoring applications, including search and rescue, security surveillance, traffic and environmental monitoring, wild animal tracking, disaster management, and patient monitoring. For example, WMSNs have recently acquired particular interest for their possible use to monitor the elderly aging at home [2]. In fact, processing video with sensor networks eliminate the need for elderly to remember to wear cumbersome instrumentation and sensors. A sensor network for health-care environment uses video transmission as sensory modality for identifying patients' behavior [2]. Another example is multimedia surveillance systems, where many different media streams (audio, video, images, textual data, sensor signals, etc.) are concerned to provide an automatic analysis of controlled environment and real-time interpretation of the scene. Video and audio sensors are utilized as multimedia facilities to enhance and complement existing surveillance systems against crime and terrorist attacks. Dependable and large-scale video and audio sensor networks, to some extent, extend the ability of law enforcement agencies to monitor areas, public events, private properties and borders [1]. Moreover, in some monitoring and security applications, non-real time audio-visual information is also beneficial to infer and record potentially relevant activities (thefts, car accidents, traffic violations) to make reports available for future query [1]. There are several other fields where WMSNs have presented unprecedented potential for applications requiring ubiquitous access to both real-time and non real-time data.

Multimedia data, including audio, video, images and scalar data, and their transmission in real-time and non-real time may merit different QoS metrics based on the application. Therefore, in addition to being power efficient, routing techniques designed for WMSNs should have the ability to facilitate application-specific service guarantee. Routing protocols designed for the traditional sensor network paradigm focus only on power consumption with the assumption that data traffic has no or loose QoS needs. As a consequence, exertion of such protocols to transmit multimedia traffic under the same limited resource and computational capacity drain the power quickly and result in operational breakdown. So, driven by acute necessity of providing QoS, routing techniques are required to be improved or re-invented. Again, energy limitation is the primal index which cannot be ignored at all.

Extensive research works on routing algorithms, protocols, and techniques have been done in the last few years to deliver multimedia content over large-scale networks, such as single-hop wired LANs (e.g., [3], [4]), multi-hop wired LANs (e.g., [5]), ATM (e.g., [6], [7]), mobile ad-hoc networks(e.g., [8], [9]) and the Internet (e.g., [10]–[14]). Routing in WSNs is very challenging due to several characteristics that are distinguished from contemporary communication and wireless ad-hoc networks [1]. As the application demand of WMSNs are escalating day by day, researchers are paying more concentration to deliver application level QoS, and striving to map these requirements to network layer metrics, such as latency, jitter, energy efficiency, reliability, packet loss, high throughput, etc. As a result, many routing techniques applicable for WMSNs have been proposed and exist in the current literature.

Although plethora of works on surveying architecture, applications and communication protocols for wireless sensor networks are performed beforehand [1], [15]–[20], this survey is distinguished from the previous efforts in that it pioneers to investigate those routing techniques which address application-specific QoS as well as energy efficiency for WMSNs. It is a thorough survey focused mainly on the network layer, describing and categorizing different power-aware (mainly real-time) routing approaches proposed in the recent years for WMSNs.

The remainder of the paper is organized as follows. We present design challenges and requirements of routing techniques for WMSNs in Section II. In Section III, existing routing approaches for traditional sensor networks are discussed along with their limitations to work for WMSNs. Then, in Section IV, current state-of-the-art of energy-efficient routing techniques for WMSNs is surveyed and categorized with a discussion on the advantages, limitations, and performance issues of each technique. A comparative summary of the current routing techniques is also provided in a tabular form. In section V, some open issues and research directions are identified. Finally, concluding remarks are given in section VI.

II. Design Requirements and Challenges

The unique requirements of WMSNs give rise to new challenges to the design of routing protocols for WMSNs. In this section, we present the most dominant and challenging factors that need be addressed in order to achieve effective communication in WMSNs.

A. QoS requirements

Different multimedia applications may have different QoS requirements, such as bounded latency or delay, bandwidth, jitter, and reliability. In the following, we describe these various QoS requirements elaborately:

- Latency: In some applications, WMSNs need to ensure stringent deadline (i.e., a physical event must be reported within a certain period of time). For example, many applications, such as intruder tracking, wild-fire monitoring, medical care and structural health diagnosis, are extremely time critical. In intruder tracking, surveillance may require the position of an intruder be reported to a command center within 15 sec so that appropriate actions can be taken in time [21]. Timeliness can be provided either in a guaranteed or a best-effort basis depending on the tolerance level of applications. Besides, end-to-end delay guarantee can be categorized into two classes: deterministic (or hard-latency bounded) and predictive (or soft-latency bounded). In hard-latency bounded systems, service that cannot be provided within its deadline is considered as a failure of the whole system. Whereas, in soft-latency bounded systems, probabilistic guarantee only is required; i.e., typically, some delay of a fraction of traffic can be tolerated. More importantly, in some applications, data in the same system may have different deadline requirements. For example, location sensory data for a fast moving target has shorter deadline than that for a slow moving target [22]. End-to-end delay measurements are facilitated through timestamps across a priori synchronized network environment, and packet forwarding over least-delay paths is preferred to maintain different deadlines requirement. Priority schemes can also be applied to differentiate among different classes of data.
- Reliability: It is defined in terms of the ability to deliver data to the destination with minimum packet loss. For example, applications, such as forest fire detection, may require that packets reach the destination or monitoring station without any loss. Again, based on the content of sensed data, different reliability constraint is needed to be imposed. For example, in fire-monitoring applications, temperature information about the regions which have normal temperatures can endure a certain percentage of loss. On the other hand, sensor data containing information about the regions which are experiencing abnormally high temperatures should be delivered to the control center with a high probability of success since it can be a sign of fire [22]. To assure such a lossless data transaction, prioritized forwarding or multipath routing can be adopted. Sending copies of the same packet over different paths increases the probability that at least one of the copies reaches the base station correctly [20]. Since unreliability of the wireless link is largely due to interference and congestion, reliability metric is often considered as reciprocal to the packet loss rate metric. Some applications of WMSNs demand assurances in terms of both timeliness and reliability. For example, in monitoring of a volcanic eruption, toxic gases or a wild fire, intruder or enemy detection, and detection of the location of survivors for rescue services, fast and reliable delivery is obligatory since late or failed delivery may result in severe disasters. Therefore, appropriate routing protocols should be carefully designed to fulfill the demand of these sophisticated applications.
- Jitter: Typically for real-time multimedia traffic, each packet is assumed to have an expected jitter requirement. Jitter is defined as the accepted variability of delay in between received packets. The presence of jitter in multimedia transmission can cause glitches, discontinuity and errors in video and audio data which is not acceptable in some applications (such as, monitoring, detection or surveillance etc.) where timely and accurate delivery of information is necessary. Therefore, synchronization should be done at the receiving point. One way to handle jitter is to buffer complete data streams prior to presentations.
- Bandwidth: Multimedia traffic by its nature requires high amounts of bandwidth. For example, the size of a monochrome uncompressed video frame in a QCIF format (176x144) is approximately 25 Kbytes [23]. In addition to the transmission of their own data, nodes may work as relay nodes due to the intrinsic, low range, multi-hop communication paradigm of WSN [23]. If a large chunk of video data is sent using single path, it can exhaust the path and result in early failure. To address the aforementioned problem, the available bandwidth can, for example, be utilized using both multiple paths and multiple channels in a spatially overlapped manner to meet high bandwidths. Therefore, high bandwidth demands should be taken into account when designing routing protocols for WMSNs.

Designing effective routing techniques that meet the above QoS requirements for different classes of traffic remains a major challenge in WMSNs.

B. Energy efficiency

Similar to WSNs, energy consumption is a prime concern in WMSNs. In some applications (e.g., environmental and habitat monitoring), WMSNs are deployed in remote and inaccessible regions (mountains, deserts, forests and rural areas) to collect multimedia information for a prolonged duration. Being the same irreplaceable, battery-operated devices, sensors in WMSNs usually consume even more energy than in WSNs. This is because multimedia applications generate high volumes of traffic, which not only require high transmission rates, but also extensive processing. While the energy consumption of traditional sensor nodes is known to be dominated by the communication functionalities, this may not necessarily be true in WMSNs. Therefore, energy-aware routing protocols for WMSNs should be designed while accounting for all forms of energy consumption.

C. Architectural issues

Since the network architecture has a great impact on the performance of routing protocols, there are some architectural issues that can serve as guidelines while designing routing protocols for WMSNs.

- Network dynamics: Some applications may require that nodes cope with their own mobility, the mobility of the sink, or the mobility of the event to be sensed (for instance, events in target detection/tracking applications are dynamic, whereas those in forest-fire monitoring are static). Taking mobility into account can be very challenging for designing routing protocols for WMSNs to deliver messages from or to moving nodes. Node movement information can impose extra burden to routing, which may cause frequent route discovery initiations. Routing protocols should then be scalable (network should be easily extendable) and flexible to topological changes.
- Data delivery models: Based on the data dissemination strategy, routing protocols in WMSNs can be divided into the following three categories: continuous, event-driven, and query-driven. The continuous delivery model requires incessant transmission of sensory data gathered at a specified rate independent of an event or a user query emanating from the sink. In the event-driven and query-driven models, the transmission of data is triggered when an event occurs or a query is generated by the sink. The continuous delivery model is not suitable for data transmission in WMSNs as the continuous compression and transfer of multimedia data is a power consuming task, and hence, may immediately result in energy depletion [23]. Therefore, practical routing protocols for WMSNs should be based on either the query-driven or the event-driven model.
- Architectural configuration: WMSNs can be consisted of either identical sensors with equal capacity in terms of sensing, computation, communication, and power (single-layered homogeneous) or special purpose sensors to perform more comprehensive tasks with varying sensing, processing and computational capabilities and energy requirement (single-layered or multi-layered, clustered). The existence of a heterogeneous set of sensors raises many technical issues related to data routing. For example, some applications might require a diverse mixture of sensors for monitoring temperature, pressure, and humidity of the surrounding environment, detecting motion via acoustic signatures, and capturing images or video tracking of moving objects [24]. In heterogeneous networks, typically data reading and reporting are generated from these sensors at different rates subject to diverse and more complex QoS constraints and can follow multiple data reporting models. Hence, routing techniques for WMSNs should consider the impact of heterogeneity of nodes in terms of their energy, memory, bandwidth requirement so that optimal lifetime and quality can be achieved.
- Channel capacity: Capacity and delay attainable on each link are location dependent, and vary continuously. Multimedia data is typically bandwidth hungry, delay intolerant, and bursty in nature. Thus, routing approaches should be designed in such a way so that data can be disseminated in a balanced and energy efficient way throughout the network under dynamic channel conditions.

D. Hole detection and bypassing

Due to the high bandwidth demands and the bursty nature of multimedia streaming data, some paths in WMSNs can get exhausted (i.e., residual energy of the nodes fall below a threshold value). These scenarios are called dynamic holes. These holes may impair the performance of multimedia applications by encumbering some routing paths. Hence, new hole-bypassing routing algorithms should be designed to facilitate the streaming of multimedia data while balancing the energy usage throughout the whole network.

Although we have covered the most dominant design objectives, it is clearly impractical to design a ”one-for-all” routing approach that address all the mentioned design objectives and requirements. All the existing WMSNs routing designs and implementations focus on specific application scenarios and put emphasis on different objectives. Hence, design challenges are mainly application- and/or network-specific.

III. Routing techniques designed for non-multimedia applications

During this past decade, considerable research efforts have been paid in developing energy-efficient routing techniques for WSNs. Majority of routing protocols designed for WSNs consider energy efficiency as the main objective with the assumption that data does not have stringent QoS requirements. As a consequence, they exhibit dissatisfactory performance whenever they are applied to QoS-bounded WMSNs. In the rest of the section, we present an abbreviated overview of the most well known routing protocols in WSNs along with their limitations.

A. Data-centric protocols

This class of protocols performs data centric routing, where the end nodes and the sensors themselves are less significant than the data itself. Thus, queries are posed for specific data rather than for data from a particular sensor, and routing is performed using the knowledge that it is the aggregate or meta-data rather than any individual data item that is important.

Sensor Protocols for Information via Negotiation (SPIN) [25] is a negotiation-based information dissemination protocol suitable for WSNs. Rather than blindly broadcasting the sensed data, sensors generate meta-data descriptions in order to represent their data about an event, and advertise the meta-data using a short ADV message. If a neighbor is interested in the data, it sends back a REQ message. Finally, sensory data is then disseminated to the interested nodes upon the reception of the REQ message. The same procedure is being repeated in the neighboring region until data has been reached to the sink node. SPIN is not applicable for multimedia routing as generating metadata descriptions for multimedia data is not a realistic task on power constrained sensor nodes. Furthermore, the ADV, REQ, and DATA flooding mechanism at each node is not appropriate for QoS constrained WMSNs applications. Moreover, guaranteed end-to-end delivery of data may not be achieved as uninterested nodes may cumber the path between the source and the sink.

Shah et al. [26] proposed to use a set of suboptimal paths in order to increase network lifetime. The paths are chosen by a means of probability functions, which depend on the energy consumption of each path. The approach argues that using the minimum energy path all the time will deplete the energy of nodes on that path. Instead, one of the multiple paths is used with a certain probability so that the whole network lifetime increases. The energy metric used here captures transmission and reception costs along with the residual energy of the nodes. Even though this method takes energy into account, it does not consider end-to-end delay. Once more, the single path strategy is not applicable for WMSN as it quickly depletes energy resources. Moreover, such a single path usage hinders the ability of recovering from a node or path failure.

Yao et al. [27] propose another data-centric protocol that observes the network as a huge distributed database system. The key idea is that the Cougar approach [27] exploits in-network data aggregation to conserve more energy. The abstraction is supported through an additional query layer that lies between the network and application layers. The data aggregation is performed by a pilot node which is selected by the query plan specified by the sink. The Cougar approach is not suitable for WMSNs due to the in-network processing overhead, node synchronization, and not taking QoS requirements into consideration.

Directed diffusion [28] is another data-centric and query-driven protocol. It aims at naming all data generated by sensor nodes by attribute-value pairs. The sink initiates a request by sending out an interest, which contains timestamps and several gradient fields defined by attribute-value pairs. Each sensor node stores the interest in its interest cache. As the interests propagate throughout the network, the gradients from the source back to the sink are established. Finally, the sink reinforces one or more paths by sending the same interest on the selected paths with a higher event rate. It is noticeable that two paths are reinforced rather than a single one. After path establishment is completed, data transport starts as defined in the matching interests. In addition to route discovery mechanisms, in-network processing may be employed to aggregate data to increase efficiency [28]. Although this is an efficient routing protocol for query-based data delivery, usage of in-network processing is a drawback in terms of multimedia data transmission as it involves a large amount of processing power which may result in early network breakdown. Moreover, directed diffusion is incompatible to handle real time traffic because it is not designed to handle QoS requirements (such as timely delivery and minimum bandwidth). Hence, this method is not suited for WMSNs. Other variants of directed diffusion, such as rumor routing [29] and Gradient-Based Routing (GBR) [30] have similar drawbacks.

B. Hierarchical routing protocols

Hierarchical routing protocols minimize energy consumption by dividing nodes into clusters. In each cluster, a node with more processing power is selected as a cluster head, which aggregates the data sent by the low-powered sensor nodes. To achieve better performance, cluster-heads perform in-network processing. One drawback of this class of protocols is the increased local communication cost between sensors and the increased processing cost of multimedia information gathered and processed at a cluster-head. Generally, cluster-heads are assumed to be capable of accessing the sink directly over longer distances. This causes high energy consumption and a low quality channel which may cause quick energy depletion while transmitting high bandwidth multimedia data. Again, selection of cluster heads are performed in randomized manner which may result in non-uniform distribution of the cluster heads in the entire network which in turn yields uneven drainage rate. In all of the following protocols, a node is selected to collect data from its proximity and make direct communication with the sink.

For example, Low-Energy Adaptive Clustering Hierarchy (LEACH) [31] is a self-organizing, adaptive clustering-based protocol that uses randomized rotation of cluster-heads to evenly distribute the energy load among the sensor nodes in the network. After the end of each round of selection, the newly elected cluster head sends to each one of the rest of its cluster nodes a consequent notification. In LEACH, data collection is performed periodically which makes the protocol not suitable for those reactive multimedia applications (e.g., event detection) where periodic data transmissions are needless, thus causing ineffectual expenditure of energy.

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) [32] is considered as an optimization of LEACH algorithms. Rather than classifying nodes into clusters, the algorithm forms chain of sensor nodes. Based on this structure, each node transmits to and receives from only closest nodes of its neighbors. The node performing data aggregation forwards the data to the node that directly communicates with the sink. In each round, a greedy algorithm is used to elect one node in the chain to communicate with the sink. The weaknesses of the protocol lies in the fact that the single leader can itself become a bottleneck in the network.

Threshold sensitive Energy Efficient sensor Network (TEEN) [33] and its adaptive version (AdaPtive) Threshold sensitive Energy Efficient sensor Network (APTEEN) [34] are similar to LEACH, i.e., both designate the transmitting nodes by using threshold mechanisms.

Due mainly to the random selection procedure of cluster-heads and the assumption of directly accessing the sink by the cluster-heads over longer distances, these protocols are not suitable to commute multimedia data without further modification.

C. Location-based protocols

This class of protocols takes advantage of the location information to make routing techniques more efficient. Unlike some data-centric approaches where queries may be broadcast blindly to all the neighboring nodes, here the direct neighbors exchange information about their location derived from global positioning system (GPS) devices, infrastructure based localization systems or ad-hoc localization systems.The energy and bandwidth are conserved as nodes are not required to keep state information beyond a single hop. In WSNs with dynamic topology changes, geographic routing has fast response and can find new routes towards the final destination quickly by using only local topology information.

Geographical and Energy Aware Routing (GEAR) [35] uses energy-aware and geographically informed neighbor selection heuristics to route a packet toward the destination region. The following protocols aim to make energy efficient routing by using location information and without making any type of aggregation.

Geographic Adaptive Fidelity (GAF) [36] is mainly designed for mobile ad hoc networks, and utilizes a virtual grid for gathering and routing messages. It saves energy by turning off unused nodes without compromising any routing fidelity, and communicates in a multi-hop manner.

Minimum Energy Communication Network (MECN) [37] and its variant Small MECN (SMECN) [38] are localized algorithms that aim to compute a network with minimum energy. This is achieved by using the location information and finding relay regions that minimize the energy by utilizing low power GPS. Both MECN and SMECN can be classified as proactive routing protocols since the latest routing information is maintained in the network. Simulation results show that SMECN uses less energy and incurs less maintenance cost than MECN. Location-based protocols, except MECN [37] and SMECN [38], generally follow routing approaches similar to those developed in the other protocol classes previously overviewed in this section. In addition, location awareness provides a reduction in latency and energy consumption, which are very crucial for multimedia applications that work on energy constrained WSN. Consequently, this class of protocols is more suitable for WMSNs applications if QoS requirements of multimedia data are met.

D. Network-flow protocol

In some approaches, route setup is modeled and solved as a network flow problem. Chang et al. [39] presented a solution to the problem of routing in sensor networks using the network flow approach. The main objective is to maximize the network lifetime by carefully defining link cost as a function of node remaining energy and the required transmission energy using that link. This approach is inapplicable for WMSNs because longer paths can still be chosen as they have high residual and low transmission energy requirements but they may not meet the end to end deadline.

IV. ROUTING TECHNIQUES FOR WMSNs

We have already mentioned that routing techniques designed for WSNs are not applicable for WMSNs as multimedia traffic is more constrained than non-multimedia traffic. Many energy-aware routing protocols have been proposed and evaluated for WMSNs in the past few years, and in this section, we survey the energy-efficient routing approaches present in the current literature. Although routing techniques for WMSNs can be categorized similarly to those designed for WSNs, we try to describe them under the light of different taxonomies to provide a comprehensive view of the contributions of the scientific community in this field and to accelerate their future convergence. According to the current research trend, the routing protocols are classified mainly according to the type and number of QoS constraints they consider. Again, we describe the routing approaches based on type of data they handle (still vs. streaming data), type of data delivery model they use (query driven, even driven), classes of algorithms (genetic algorithm, supervised learning, clustered-control algorithm etc.) they adopt, and the hole-bypassing approaches they use.

A broad classification of different routing techniques for WMSNs is shown in Fig. 2 where the numbers in future indicate the references.

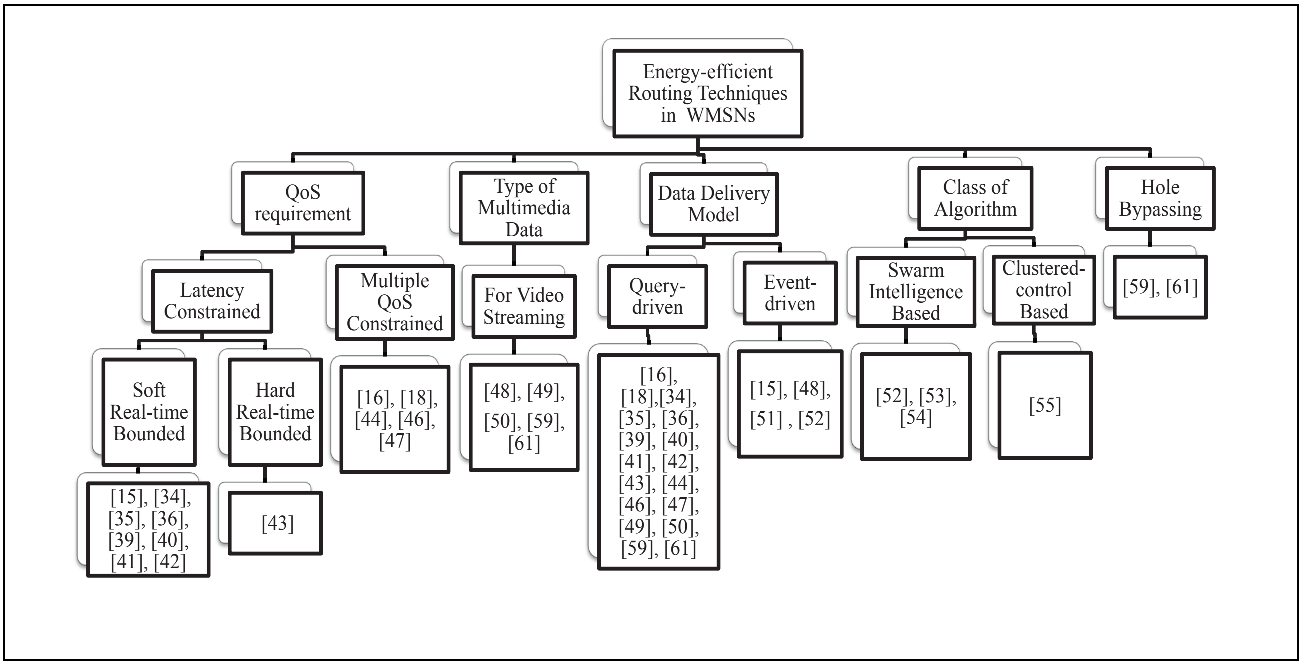

Fig. 2: Classification of Energy-efficient Routing Techniques for WMSNs

A. Latency constrained routing

In this section, we survey protocols based on the type of end-to-end delay guarantee (soft or hard real time bounded latency) they provide.

1) Protocols providing soft real time delay guarantee: Sequential Assignment Routing (SAR) [40] is the first routing protocol considering QoS (e.g., end-to-end delay) and energy efficiency for sensor networks which includes a table driven multi-path routing and path restoration technique to create trees routed from one-hop neighbor of the sink. The objective of the SAR routing algorithm is to minimize the average weighted QoS metric throughout the lifetime of the network. The multipath routing scheme ensures fault-tolerance, and path restoration technique eases the recovery in case of node failure. The limitation of SAR approach is that it suffers from the overhead of maintaining the tables and status information for each sensor node when number of nodes is huge.

RAP [21] is a soft real time delay bounded, location-aware and priority-based routing protocol which pioneers in considering deadline issues in multi-hop wireless multimedia sensor networks. RAP provides convenient mechanisms to be used in both query-initiated and event-initiated applications. RAP implements a differentiated priority-based Velocity Monotonic Scheduling (VMS) policy, which is based on a notion of packet requested velocity suitable for packet scheduling in sensor networks. Each packet is expected to make its end-to-end deadline if it can move toward the destination at its requested velocity, which reflects its local urgency. Compared with non-prioritized packet scheduling, VMS improves the deadline miss ratios of sensor networks by giving higher priority to packets with higher requested velocities. Also, VMS can perform better than deadline-based packet scheduling because velocity reflects the local urgency at each hop more accurately when packets with the same deadline have different distances to their destinations. Since the requested velocity can be locally determined assuming that each sensor node knows its own location, a combination of VMS and geographic forwarding (GF) provides a localized and scalable protocol for real-time multimedia communication on sensor networks. But for multimedia data, such greedy forwarding can deplete the bandwidth quickly. The limitation of RAP is that it does not consider any alternate approach for such problem. Again, while calculating priority, the scheme does not consider the number of hops that the packet has to traverse in deciding the priority. Most importantly, there is no direct metric to show how energy is conserved in multimedia routing.

Energy Aware QoS (EAQoS) [41] is another early proposed routing protocols that can run efficiently along with best-effort traffic. It aims to discover an optimal path in terms of energy consumption and error rate along which the end-to-end delay requirement can be satisfied. The proposed protocol finds a least-cost and delay-constrained path for real-time data. The cost of the link is defined to capture communication parameters, such as the residual energy in the nodes, the transmission energy, and the error rate. Moreover, it maximizes the throughput for non-real-time data by adjusting the service rate for both real-time and non-real-time data coexisting at sensor nodes. In order to provide both real-time and non-real time best-effort traffic simultaneously, a class-based queuing model is used. The protocol is based on a two-step strategy incorporating both link-based costs and end-to-end constraints as follows:

- First, some k-least cost paths are calculated by using an extended version of Dijkstra's algorithm without considering the end-to-end delay.
- Then, among all the candidate paths that meet the end-to-end real-time QoS requirements, the one that maximizes the throughput for non-real-time traffic is chosen.

EAQoS consistently performs well with respect to real-time and energy metrics, but the main drawback of this protocol is that it does not use any priority scheme to account for the different end-to-end delay requirements that real-time multimedia traffic may have. So, this technique is not suitable for real time streaming applications that may coexist along with real time applications. Another noticeable drawback is that, when calculating the end-to-end delay the scheme does not consider several network delays such as MAC-related channel access delays, or actual packet queuing delay at intermediate relaying nodes. The consideration of only propagation delay and average queuing delay in calculating end-to-end delay limits the ability of the protocol to satisfy the actual QoS needs. In addition, the bandwidth ratio is initially set the same for all the nodes, which does not provide adaptive bandwidth sharing for different links. Moreover, the algorithm requires complete knowledge of the network topology at each node in order to calculate multiple paths, thus limiting the scalability of the approach.

SPEED [42] is proposed to support real time traffic with delay requirements while considering prioritization. SPEED is a spatio-temporal, priority-based, QoS-aware routing protocol for sensor networks that provides soft real-time, end-to-end delay guarantees. The protocol requires that each node maintain localized information with minimal control overhead, and uses non-deterministic geographic forwarding to find paths. The main objective of this work is to support a spatio-temporal communication service with a given maximum delivery speed across the network. The routing module in SPEED is called Stateless Geographic Non-Deterministic forwarding (SNFG), and works with other four modules at the network layer. They are:

- Beacon Exchange: used to collect location information about nodes.
- Delay Estimation: made by calculating the elapsed time when an acknowledgement (ACK) is received from a neighbor as a response to a transmitted data packet. By looking at the delay values, SNGF selects the nodes which meet the speed requirement.
- Neighborhood Feedback Loop (NFL): responsible for providing the relay ratio which is calculated by looking at the miss ratios of the neighbors of a node (the nodes which could not provide the desired speed), and is fed to the SNGF module. If the relay ratio is less than a randomly generated number between 0 and 1, the packet is dropped.
- Back-pressure Rerouting: used to handle holes, and mitigate congestion by informing the nodes up to the source to search for alternate route.

SPEED performs better than Dynamic Source Routing (DSR) [43] and Ad-hoc On-Demand Vector Routing (AODV) [44] in terms of end-to-end delays and miss ratios. Moreover, the total transmission energy is less because of minimum control packet overhead, and a network wide load balancing is achieved by the even distribution of traffic. One drawback of SPEED is that energy consumption is not addressed directly in the protocol. For more realistic understanding of SPEED's energy consumption, there is a need for comparing it to an energy-aware routing protocol. In addition, SPEED does not adopt differentiated packet prioritization, and each forwarding node can only forward the packet at some speed less than or equal to the maximum achievable speed. However, using SPEED protocol it is not possible to forward a packet at a higher speed, even if the network can support it. This is not an optimal behavior whenever we are dealing with real time streaming multimedia data which may contain data with different importance levels. Due to the highly dynamic link and route characteristics, the idea of per-flow reservation appears to be non-scalable, and hence, SPEED might present some scalability issues when dealing with large WSNs.

Real-time Power-Aware Routing (RPAR) [45] protocol pioneers the approach of incorporating energy efficiency in real-time communication. RPAR achieves application specific end-to-end delay guarantee at low power by dynamically adjusting transmission power and routing decisions based on the workload and packet deadlines. Another distinctive feature of RPAR is that it calculates average link quality taking link variability into consideration. RPAR also handles realistic and dynamic properties of WMSNs such as lossy links, limited memory, and bandwidth. The usage of localized information adds scalability feature with RPAR. In addition, the novel neighborhood management mechanism apart from periodic beaconing scheme adopted by SPEED [42] and MMSPEED [22] is used to maintain and discover neighboring nodes. The unique forwarding policy and neighborhood management of RPAR together can introduce significant power reduction hence results in prolonged network lifetime with desired real-time guarantee. The drawback of the protocol is that it shows degraded performance in handling large hole and sudden congestion.

Pothuri et al. [46] propose a novel heuristic solution for the delay-constrained, energy-efficient routing problem in sensor networks. A hierarchical network architecture in conjunction with routing framework models to handle access delays caused by the MAC layer are the main features of the proposed protocol. The forwarding policy works as follows: Given a delay bound, the task is to find a path from a sensor node to the sink with the lowest energy consumption, such that the total transfer delay incurred along the path is less than the delay bound. However, the assumption that each sensor node can reach the sink directly using its long-range radio is highly undesirable for WMSNs as it may cause quick energy depletion.

Yuan et al. [47] propose an energy-efficient real-time routing protocol using a concept similar to that of SPEED. The proposed protocol adopts an effective transmission (ET) concept that ensures that the relaying candidates are not only closer to the sink, but also farther from the source node with respect to its preceding node. In addition, the path's end-to-end delay guarantee is expressed in terms of the sum of point-to-point Constrained Equivalent Delay (CED). Intermediate nodes independently of one another decide on their next forwarding nodes according to the value of the CED of the link.

Another robust, energy-aware routing technique for WMSNs that ensures timeliness is proposed by Khalid et al. [48]. The approach is based on a Logical Network Abridgement (LNA) technique, which is capable of measuring the underlying path diversity, intrinsic network resiliency to congestion, failures, and attacks. The protocol considers separate time aware and energy aware cost functions. It, however, exhibits poor scalability, because it tends to maintain information for the whole network. Future research is needed on the selection of parameter values, and understanding the relationship between different parameters in the cost functions and making the routing more adaptive.

2) Protocols providing hard real time delay guarantee: Ergen et al. [49] propose a power efficient routing algorithm that ensures hard delay guarantee for WMSNs. It adopts a single sink model and aims at maximizing the lifetime of a WMSN by adjusting the number of packets traversing throughout the network. To achieve this goal, first the delay constraint is excluded and lifetime maximization is formulated as a linear programming (LP) problem, and solution is implemented in a distributed manner which uses an iterative algorithm to approximate the centralized optimal one. Then, delay guarantee is ensured into the energy efficient routing by limiting the length of routing paths from each node to the sink. The protocol can guarantee end-to-end delay requirement and prolong network lifetime. However, it is noticeable that it is not flexible enough to meet application-specified delay bound generally.

B. Multiple QoS-constrained routing

Multi-path and Multi-SPEED (MMSPEED) routing protocol [22] supports probabilistic QoS guarantee by provisioning QoS in two domains, timeliness and reliability. MMSPEED adopts a differentiated priority packet delivery mechanism in which QoS differentiation in timeliness is achieved by providing multiple network-wide packet delivery speed guarantees. The scheme employs fully distributed localized geographic packet forwarding augmented with dynamic compensation, which compensates for local decision inaccuracies while a packet travels towards its destination. MMSPEED needs the support of IEEE 802.11e at the MAC layer with its inherent prioritization mechanism based on the Differentiated Inter-Frame Spacing (DIFS). Each speed level is mapped onto a MAC layer priority class. For supporting service reliability, probabilistic multi-path forwarding is used to control the number of delivery paths based on the required end-to-end reaching probability. In the scheme, each node in the network calculates the possible reliable forwarding probability value of each of its neighbors to a destination by using the packet loss rate at the MAC layer. According to the required reliable probability of a packet, each node can forward multiple copies of packets to a group of selected neighbors in the forwarding neighbor set to achieve the desired level of reliability. MMSPEED could use its redundant path selection scheme for load balancing, which is not only for reliability enhancement, but also to improve the overall network lifetime. Although MMSPEED considers energy consumption indirectly by not using unnecessary flooding, it does not pay heed to an individual node's energy situation. Moreover, it does not consider the number of hops that the packet has to traverse in deciding priority, where the number of hops is a more pragmatic measure for priority assignment than the distance between source and destination.

Hamid et al. [50] propose a timeliness and reliability QoS-aware, localized, multi-path and multi-channel protocol, where routing decision is made according to the dynamic adjustment of the required bandwidth and path-length-based proportional delay differentiation for real-time data. The proposed protocol works in a distributed manner to ensure bandwidth and end-to-end delay requirements of real time data. At the same time, the throughput of non-real time data is maximized by adjusting the non-real time data. A differentiated priority based classifier and scheduler which schedules packets according to delay and bandwidth requirement are the power of this QoS-aware approach. The delay bound is calculated based on propagation delay, transmission delay and the switching delay and the hop count is used for dynamic bandwidth adjustment to boost up the time critical real-time traffic so that they can meet the deadline. The proposed protocol experiences less average delay for real time packets and maximizes the throughput of non-real time traffic exploiting multiple paths. Also, the processing hubs reduce redundant data and channel assignment technique results in less collisions. The limitation of the protocol is that it does not adopt any alternative mechanism to handle the delay whenever the buffer size increases and switching overhead affects the performance of the protocol. Like SPEED [42] and MMSPEED [22], no direct metric has been used to assess how efficiently it manages power consumption.

Distributed aggregate routing (DARA) [51] is a multi-sink, multipath, and location-aware protocol to allow the real-time constrained data packets to pass through the shortest route and the non-time constrained traffics to follow a longer route keeping the shortest route free for time constrained packets. DARA algorithm discovers the most suitable relaying nodes towards each sink for both the time constrained and non-time constrained packets. To ensure reliability, only the source node sends multiple copies of a packet (packet duplication), if required, towards multiple sinks while the intermediate nodes forward the received copy only towards the destination. The number of duplicate packets at each source node is calculated by solving an optimization problem. Forwarder nodes use class based packet scheduling based on the delay-deadline values. The protocol achieves energy efficiency by implementing power controlled transmission and reduced number of retransmission. Furthermore, DARA reduces the standard deviation of energy consumption by using residual energy of target node as routing metric, which in turn increases the network lifetime. Again, it optimizes the tradeoffs between the reliability and the delay guarantee while improving the spatial balance of energy burdens. DARA effectively improves the reliability, delay guarantee and energy efficiency.

Real Time routing protocol with Load Distribution (RTLD) [52] provides efficient power consumption and high packet delivery ratio in WMSN. The major advantage of RTLD is that it can deliver packets within their end-to-end deadlines, while minimizing the network miss ratio and power consumption. It combines the geocast forwarding with link quality, maximum velocity and remaining power to achieve the real time routing in WMSN. The remaining power is used to assist for mitigating routing holes problem due to power expiration. It also has a novel neighborhood selection and power management policy. RTLD shows better performance than AODV [44], DSR [43] and RPAR [45] in terms of average end-to-end deadline, network miss ratio, and network lifetime.

Sen and Ukil [53] propose a query-based, adaptable, energy-efficient routing protocol for WMSNs that can assure multiple QoS requirements such as reliability and latency. For ensuring path reliability, the algorithm constructs multiple alternate paths from source nodes to the sink nodes, and for guaranteeing data reliability, it sends multiple copies of the same message. The latency is minimized by allowing the nodes to transmit with more power.

Mahapatra et al. [54] propose an energy aware dual-path routing scheme for real-time traffic, which balances node energy consumption to prolong the network lifetime, considers network congestion to reduce network-wide routing delay and enhances the reliability of the packets reaching the destination by introducing minimal data redundancy.

C. Routing for video streaming

In WMSNs, real time video streaming data generally has a soft deadline, calling then for shortest path routing approaches with the minimum end-to-end delay. Also, transmission requirements in terms of bandwidth can be several times higher than the maximum transmission capacity (bandwidth) of sensor nodes, needing then multipath routing. Some protocols are proposed to explicitly handle real time streaming by taking both end-to-end latency and bandwidth into consideration.

Optimized energy-delay sub-network routing (OEDSR) [55] is a cluster-based event-driven multihop energy efficient approach addressing the end-to-end delay constraint. The fully distributed OEDSR protocol calculates the available energy, average end-to-end latency values of the links and the distance from the sink to determine the best next-hop forwarding node. The protocol ensures that the selected path from the cluster-head to sink to be loop-free, power-efficient and has the least end-to-end delay. In addition, the lifetime of the network is maximized since the energy is taken into account while selecting nodes from a route. Moreover, when a node loses more energy an alternate path is computed to maintain the load of the network. OEDSR achieves lower average end-to-end delay, fewer collisions and less energy consumption than DSR [43] and AODV [44].

Directional Geographical Routing (DGR) [56] protocol, which investigates H.26L real-time video communications in WMSNs, where video streams are transmitted under a number of resource and performance constraints, such as bandwidth, energy and delay. DGR divides a single video stream into multiple sub-streams and exploits multiple disjoint paths to transmit these sub-streams in parallel in order to make the best of limited bandwidth and energy in WSNs and to achieve a reliable delivery. Simulation results show that DGR exhibits high delivery ratio and low end-to-end delay. However, it tackles path failures with local repairs at the cost of additional overheads and transmission latency. In addition, DGR assumes that any node can send video packets to the sink at any instance, which limits its practicality of deploying it for large networks.

Poltis et al. [57] propose a power efficient multipath video packet scheduling scheme for minimum video distortion transmission over WMSNs. A modified LEACH (hierarchical) protocol has been used to select multiple paths for the transmission of video packets which improves the aggregate data rate of the network and minimizes the traffic load handled by each node. To support requested video source data rates in lossy channel conditions, two scheduling algorithms have been proposed to adapt the source requirements to the channel capacity by dropping less important video packets. The proposed power aware video packet scheduling can achieve energy efficiency in the wireless multimedia sensor network by minimizing the power dissipation across all the nodes, while the perceived video quality is kept to very high levels even at extreme network conditions when many sensor nodes dropped due to power consumption and high background noise in the channel.

D. Data-delivery model based routing

Most of the aforementioned protocols are query-initiated. Real-time and Energy Aware Routing (REAR) [58] is an event-driven protocol which uses metadata to establish multi-path routing for reducing the energy consumption. A cost function is constructed to evaluate the consumption of bandwidth on the transmission links which trades off the relationship of energy and delay, and then QoS routing is chosen. The end-to-end delay of multi-hop routing not only depends on transmission distance, but also relies on relay nodes processing and queuing delay. Because of the high bandwidth requirement of real time multimedia data, a classified queue model is introduced at each node to deal with both real time and non-real time data. This protocol saves energy by means of activating image sensors when monitoring events occur and using metadata instead of real data in routing setup. But in case of streaming applications, the idea of meta-data is not a good choice as the meta-data for streaming data can itself be very huge and result in high energy and bandwidth consumption.

E. Routing approaches addressing different classes of algorithms

In this section, routing approaches based on different classes of algorithm are presented.

1) Swarm intelligence based routing: Recently, different adaptive classes of algorithms such as ant-colony based and genetic algorithms are being considered to take intelligent routing decisions in WMSNs. As a reflection, many routing techniques based on such algorithms are evident in the current literature.

Ant-based Service Aware Routing (ASAR) [59] is a hierarchical protocol that incorporates reinforcement learning to routing. ASAR periodically chooses three suitable paths to meet diverse QoS requirements from different kind of services (event-driven, data query and stream query) by positive feedback mechanism used in ant-based algorithms, thus maximizing network utilization and improving network performance. It maintains optimal path table and pheromone path table at each cluster head. Routing selection for different data services is made based on delay, packet loss rate, bandwidth and energy consumption required by the type of traffic. Besides the bottleneck problem of hierarchical models, new optimal path setup due to congestion requires extra calculation which may decrease network performance by engaging extra energy for large networks.

Peng et al. [60] propose an adaptive QoS and energy-aware routing approach that uses an improved biological ant colony algorithm for WMSNs not only to meet QoS requirements in an energy-aware fashion, but also balance the node energy utilization to maximize the network lifetime. The special forwarding ants are used to discover optimum routing paths between the sensor nodes and the sink nodes in terms of distance, delay, packet loss, bandwidth, and energy levels. The simulation results provide evidence that the proposed algorithm has good performances in different WSN scenarios. Extensive performance analysis under different network load is yet to be performed.

Zongwu et al. [61] propose a novel genetic algorithm based on game theory. The forwarding policy is based on satisfying the necessary QoS parameters (such as, guaranteed bandwidth and end-to-end delay, minimum cost and maximizing network lifetime simultaneously). In the proposed algorithm, mixed strategy Nash equilibrium of crossover game instead of probability crossover has been used. The crossover game in the routing problems is based on the assumption that each node has restricted energy, and each node is inclined to get maximal whole network benefit but pay out minimum cost. Moreover, individual routes are treated as the players of crossover game.

2) Clustered-control based routing: As power control, proper election of cluster-head and avoidance of unnecessary communication links for the whole network are more crucial in WMSNs, routing based on clustered-control algorithms has become an active research area recently.

Haiping and Ruchuan [62] propose a novel clustered-control algorithm based on location information, energy, priority of coverage and multi-layered architecture, which is different from connection prediction scheme [63] and two-hop clustered image transmission scheme [64]. This approach selects a cluster-head according to geographical locations and remained energy at the nodes and ensures the higher coverage rate for the cluster-head by a priority mechanism to avoid the concentrated and marginal distribution of cluster-heads. This approach reduces the energy cost by increasing the sleeping nodes during non-media data transmission phase and adding many intermediate nodes to forward data during multimedia media data transmission which in turn prolongs the lifetime of the network. Simulation results show that this protocol achieves prolonged service time and decreased energy consumption than GAF [36] and ASCENT [65]. The limitation of this approach is lack of showing experimental results in claiming that it improves reliability, load-balancing, uniform distribution of cluster-heads and strengthens the backbone routing paths in the networks.

F. Hole-bypassing routing

Two Phase Geographical Greedy Forwarding (TPGF) [66] protocol focuses on exploring the maximum number of approximately optimal node-disjoint routing paths in the network layer in order to minimize the path lengths and the end-to-end transmission delay while taking the energy constraints into consideration. TPGF protocol includes two phases: The first phase explores all possible routing paths, whereas the second phase optimizes the found routing paths to shorten transmission distance. The routing algorithm finds one path per execution and can be executed repeatedly to find more node-disjoint routing paths. Hole-bypassing is considered as the primary design goal of the proposed method as holes can impede the performance of multimedia streaming data transmission. Simulation results show that using this algorithm, both static and dynamic holes can be bypassed efficiently [66]. When compared to Greedy Perimeter Stateless Routing (GPSR) [67], TGPF shows better performance. GPSR is not always able to find the routing path when it exists in the network. It stops exploring when no neighbor node is available for the next hop transmission which is called a block situation. TGPF effectively handles such situations using an approach called Step Back and Mark approach where a sensor node steps back to its previous hop node and marks itself as a block node when there is no neighbor node available for the next-hop transmission. The Step Back and Mark is repeatedly executed until a sensor node successfully finds a routing path to the base station. This approach guarantees that a path is found if it exists in the network. It has also been shown that the notorious local minimum problem, which exists in GPSR, is not present when TGPF is used. Multipath routing has also been achieved by showing the capability of finding additional routing paths for increasing the multimedia streaming data transmission when it is necessary. The drawback of TPGF is that it needs to build a complete map of the network topology to select the optimal routing path between the source and destination which limits its adaptability in large-scale, high density and frequent mobility situations.

Another routing protocol which mitigates the dynamic hole problems in WMSNs efficiently is Geographic Energy-Aware Multipath Stream-based protocol (GEAMS) [68]. GEAMS is a geographical, multipath localized routing protocol designed to handle multimedia streaming data by maintaining both QoS restriction and energy efficiency. GEAMS routes information based on GPSR [67] functionalities while maintaining local-knowledge for delivering this information on multipath basis. In GPSR, same nodes in close vicinity with the sinks are chosen repeatedly which may cause early failure of most of the nodes. Unlike that, in GEAMS routing protocol, data streams are routed by different nodes and decisions are made locally at each hop. Decision policy at each node is based on:

- The remaining energy at each neighbor
- The number of hops made by the packet before it arrives at this node
- The actual distance between the node and its neighbors, and
- The history of the packets forwarded belonging to the same stream.

GEAMS has two modes, Smart Greedy Forwarding and Walking Back Forwarding. The first mode is used when there is always a neighbor closer to the sink node than the forwarder node. The second is used to avoid and bypass holes. In addition, to meet the multimedia transmission constraints and to maximize the network lifetime, GEAMS also exploits the multipath capabilities of the WSN to make load balancing among nodes. GEAMS is more suitable for WMSNs than GPSR as it ensures uniform energy consumption and meets the delay and packet loss constraint.

In Table I, we compare and summarize the aforementioned routing protocols for WMSNs based on types of QoS constraints, multimedia data they can handle, data delivery models, network architecture, and hole-bypassing techniques.

V. Open research issues

Due to the persistent proliferation of novel QoS critical applications, researchers are continually facing new challenges for improvising energy efficiency with stringent QoS guarantees in designing routing approaches for WMSNs. Despite the large volume of research activities and the rapid and significant progress are made in the recent years, routing in WMSNs still harbor many open issues which are still to be resolved. In this section, a broad range of research issues is outlined for future investigation.

- Energy efficiency: There is no doubt that energy efficiency should get the foremost importance in designing routing protocols for sensor networks. But for WMSNs, the QoS requirements must have equal importance. So, there is always a trade-off between multimedia service guarantee and energy efficiency. For example, if we consider the case of energy efficiency and end-to-end delay, by a larger transmission power with high energy consumption, a message delivery velocity can be increased which can effectively decline the end-to-end delay, but results in shorter lifetime. Finding optimal solution to balance energy efficiency and QoS requirements trade-off is an everlasting area of study until auto-rechargeable sensor devices are invented. A reasonable joint optimization of tunable metrics should be defined to measure the performance of designed routing protocols.
- Multi-constrained QoS guarantee: While offering QoS support for multimedia traffic, routing algorithms should be flexible to support different application-specific QoS requirements (such as energy efficiency, end-to-end delay, reliability, delay jitter, bandwidth consumption) in the heterogeneous traffic environment [19]. Real-time delay guarantee has been the main QoS requirement considered in WMSNs so far. However, diverse reliability constraint also needs to be considered. In addition, meeting delay jitter constraint is another intricate problem in QoS routing [41]. So, designing flexible integrated architecture with adaptive cost functions to provide multi-constrained QoS guarantee can be reckoned as an interesting area of future research.
- New class of algorithms: Routing in sensor networks maintains neighbor information on neighbors' states

TABLE I

COMPARISON OF ROUTING PROTOCOLS FOR WMSNS

Routing protocol	Architecture		Location awareness	Multipath capability	Energy efficiency	Bounded latency	Reliable delivery	Data-delivery model		Hole bypassing
	Flat	Hierarchical						Query-driven	Event-driven	
SAR [40]	✓			✓	✓	✓		✓		
RAP [21]	✓		✓			✓		✓	✓	
EAQoS [41]		✓			✓	✓		✓		
SPEED [42]	✓		✓			✓		✓		
RPAR [45]	✓		✓		✓	✓		✓		
Pothuri et al. [46]		✓			✓	✓		✓		
Yuan et al. [47]	✓		✓		✓	✓		✓		
Khalid et al. [48]	✓			✓	✓	✓		✓		
Ergen et al. [49]	✓				✓	✓		✓		
MMSPEED [22]	✓		✓	✓		✓	✓	✓		
Hamid et al. [50]	✓			✓		✓	✓	✓		
DARA [51]	✓		✓	✓	✓	✓	✓	✓		
RTLD [52]	✓		✓	✓	✓	✓	✓	✓		✓
Sen & Ukil [53]	✓			✓	✓	✓	✓	✓		
Mahapatra et al. [54]	✓			✓	✓	✓	✓	✓		
OEDSR [55]	✓			✓	✓	✓			✓	
DGR [56]	✓			✓	✓	✓	✓			
Poltis et al. [57]		✓		✓	✓	✓	✓			
REAR [58]	✓				✓	✓			✓	
ASAR [59]		✓			✓	✓	✓	✓	✓	
Peng et al. [60]	✓				✓	✓	✓			
Zongwu et al. [61]	✓				✓	✓				
Haiping & Ruchuan [62]		✓	✓		✓		✓			
TPGF [66]	✓		✓	✓	✓	✓	✓			✓
GEAMS [68]	✓		✓	✓	✓	✓	✓			✓

and potentially many other factors in order to make informed decisions. Challenges arise in performing accurate and adaptive information discovery. For routing in WMSNs, supervised and reinforcement learning based algorithms are interesting classes of algorithms that require further explorations to devise QoS routing techniques for adaptive information discovery.

- Mobility: Almost all the WMSN routing protocols existing in the current literature do not take mobility into consideration. However, the targets, sensors and sinks, may be highly movable. Recent advances by adding mobile sensors in WSNs have improved the performance well, including coverage and energy efficiency. Furthermore, mobility of sensors and sinks may bring benefit to real-time delivery guarantee. Thus, supporting multimedia in highly dynamic WMSNs can be viewed as an interesting area for future investigation.
- Multiple sources and sinks: So far, most routing techniques designed for WMSNs consider sending data from single source to single sink. A diverse network with multiple sinks may be required to receive event information simultaneously. Devising efficient routing techniques for such area can be considered as a new area of exploration.
- Dynamic hole bypassing: In WMSNs, dynamic holes may occur at a high frequency. To overcome holes, routing may choose alternate path that is not optimal which in turn can cause extra delay. More explorations are required to adapt to topology changes caused by holes.
- Secure routing: Current routing protocols mainly focus on optimizing application specific nature of the sensor networks, but do not take security into account. But WMSNs transmit multimedia information which provides more detailed information. Such information leakage can be more malicious. In this regard, secure routing is an issue that needs further attention. Moreover, although recent protocols have not been designed with security as a goal, it is important to analyze their performance when security will be incorporated with the routing protocols.
- Cross-layer awareness: It is not possible to provide optimal QoS routing solution solely by the network layer. Routing requires the involvement of physical, data-link, and network layer in order to be effective and power efficient. Thus, when incorporated with other layer functionalities, routing can reveal as a stronger solution to many unsolved problems.
- Multi-channel access: Multi-channel routing brings great potentials for handling high bandwidth data in WMSNs. Although many solutions have been proposed in the literature to address issues pertaining to physical- and MAC-layers in multi-channel access networks [69]–[74], there is still room for devising and proposing efficient routing schemes that take advantage of multi-channel access capability to promote efficient data delivery in WMSNs.
- Integration with other networks: Most of the applications in security surveillance and environmental monitoring require the data collected from the sensor

nodes to be transmitted to a server so that further analysis and/or actions can be done and/or taken. However, while travelling to the sink, collected data or requests may traverse multiple different networks. For instance, it is possible, or even likely, that requests or collected data traverse through the Internet backbone before reaching the sink. End-to-end QoS guarantees can become more complicated because of the inherent differences in the nature of different networks. Decoupling of reliability and routing parameters at the network boundaries and seamless integration of schemes seem well suited.

VI. Conclusion

The emergence of WMSNs in supporting wide variety of applications are motivating the researchers to do more research on the routing protocols. The common objective is to provide QoS guarantee while using energy conserving strategies. In this paper, we present a comprehensive review on research challenges and the state-of-the art of energy-aware routing techniques for WMSNs, and highlight the advantages and performance issues of each routing protocol and algorithm. Finally, open issues are provided in order to stimulate more research interests in those unexplored areas. It is doubtless that being blessed by the growing advancement of hardware technology, WMSNs will reveal as a powerful technology in near future. So, developing efficient routing protocols for WMSNs appears to be a promising direction of future research.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdury, "A Survey on Wireless Multimedia Sensor Networks," *Computer Networks (Elsevier)*, vol. 51, no. 4, pp. 921–960, March 2007.

[2] T. Teixeira, E. Culurciello, E. Park, D. Lymberopoulos, A. B. Sweeney, and A. Savvides, "Address-event imagers for sensor networks:Evaluation and modeling," in *SPOTS '06*, November 2006.

[3] W. Zhao, J. Stankovic, , and K. Ramamritham, " A Window Protocol for Transmission of Time Constrained Messages," *IEEE Transactions on Computers*, vol. 39, no. 9, pp. 1186–1203, September 1990.

[4] K. Zuberi and K. G. Shin, "Design and Implementation of Efficient Message Scheduling for Controller Area Network," *IEEE Transactions on Computers*, vol. 49, no. 2, pp. 182–188, February 2000.

[5] D. Kandlur, K. G. Shin, and D. Ferrari, " Real-time communication in multi-hop networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 10, pp. 1044 – 1056, October 1994.

[6] S. K. Kweon and K. G. Shin, " Providing deterministic delay guarantees in ATM networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 6, pp. 838–850, December 1998.

[7] C. Li, R. Bettati, and W. Zhao, "Static Priority Scheduling for ATM Networks," in *the Proceedings of IEEE Real-time Systems Symposium*, December 1997.

[8] L. Chen and W. B. Heinzelman, "A Survey of Routing Protocols that Support QoS in Mobile Ad Hoc Networks," *IEEE Network*, vol. 21, no. 6, 2007.

[9] L. Hanzo-II and R. Tafazolli, "A Survey of QoS Routing Solutions for Mobile Ad Hoc Networks," *IEEE Communications Surveys and Tutorials*, vol. 9, no. 2, 2007.

[10] J. Liebeherr, D. E. Wrege, and D. Ferrari, " Exact admission control in networks with bounded delay services," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 885–901, 1996.

[11] S. Wang, D. Xuan, R. Bettati, and W. Zhao, "Providing Absolute Differentiated Services for Real-time Applications in Static-Priority Scheduling Networks," in *the Proceedings of IEEE INFOCOM*, 2001.

[12] I. Stoica and H. Zhang, "Providing Guaranteed Services Without Per Flow Management," in *the Proceedings of SIGCOMM*, 1999.

[13] O. Younis and S. Fahmy, "Constraint-based routing in the internet: Basic principles and recent research," *IEEE Communications Surveys and Tutorials*, vol. 5, no. 1, 2003.

[14] J. Soldatos, E. Vayias, and G. Kormentzas, "On the building blocks of quality of service in heterogeneous IP networks," *IEEE Communications Surveys and Tutorials*, vol. 7, no. 1, 2005.

[15] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirc, " Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, March 2002.

[16] K. Akkaya and M. Younis, " A Survey on Routing Protocols for Wireless Sensor Networks," *Ad Hoc Networks (Elsevier)*, vol. 3, no. 3, pp. 325–349, May 2005.

[17] R. Rajagopalan and P. K. Varshney, "Data-Aggregation Techniques in Sensor Networks: A Survey," *IEEE Communications Surveys and Tutorials*, vol. 8, no. 4, 2006.

[18] N. Pantazis and D. Vergados, "A survey on power control issues in wireless sensor networks," *IEEE Communications Surveys and Tutorials*, vol. 9, no. 4, 2007.

[19] Y. Li, C. S. Chen, Y. Song, and Z. Wang, "Real-time QoS support in wireless sensor networks: a survey," in *the Proceedings of (FeT'07), Toulouse, France*, November 2007.

[20] S. Misra, M. Reisslein, and G. Xue, " A survey of multimedia streaming in wireless sensor networks," *IEEE Communications Surveys and Tutorials*, vol. 10, pp. 18–39, 2008.

[21] C. Lu, B. M. Blum, T. F. Abdelzaher, J. A. Stankovic, and T. He, "RAP: A Real-time Communication Architecture for Large-Scale Wireless Sensor Networks," in *the Proceedings of the Eighth IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS' 02)*, September 2002.

[22] E. Felemban, C. Lee, and E. Ekici, "MMSPEED: Multipath multi-SPEED protocol for QoS guarantee of reliability and timeliness in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 738–754, 2006.

[23] E. Gurses and O. B. Akan, "Multimedia Communication in Wireless Sensor Network," *Annales des Telecommunications*, vol. 60, no. 7-8, pp. 872–900, 2005.

[24] J. N. AL-Karaki and A. E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey," *IEEE Wireless Communications*, vol. 11, no. 6, December 2008.

[25] W. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks," in *the Proceedings of ACM/IEEE Mobicom Conference (MobiCom '99), Seattle, WA*, August 1999.

[26] W. R. Shah and J. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks," in *the Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Orlando, FL*, March 2002.

[27] Y. Yao and J. Gehrke, " The cougar approach to in-network query processing in sensor networks," *SIGMOD Record*.

[28] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, " Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, February 2003.

[29] D. Braginsky and D. Estrin, "Rumor Routing Algorithm for Sensor Networks," in *the Proceedings of ACM Workshop on Sensor Networks and Applications (WSNA), Atlanta, GA*, October 2002.

[30] C. Schurgers and M. B. Srivastava, "Energy Efficient Routing in Wireless Sensor Networks," in *the Proceedings of IEEE MILCOM on Communication for Network-Centric Operations: Creating the information Force, McLean, VA*, 2001.

[31] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless sensor networks," in *the Proceedings of the Hawaii International Conference System Sciences, Hawaii*, January 2000.

[32] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power Efficient Gathering in Sensor Information Systems," in *the Proceedings of IEEE Aerospace Conference, Big Sky, Montana*, March 2002.

[33] A. Manjeshwar and D. P. Agrawal, "TEEN: A Protocol For Enhanced Efficiency in Wireless Sensor Networks," in *the Pro-*

ceedings of 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, San Francisco, CA, April 2001.

[34] ——, "APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks," in *the Proceedings of International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, Ft. Lauderdale, Fl*, April 2002.

[35] Y. Yu, D. Estrin, and R. Govindan, " Geographical and Energy Aware Routing: A Recursive Data Dissemination Protocol for Wireless Sensor Networks," UCLA Computer Science Department Technical Report UCLA-CSD TR-010023, Tech. Rep., May 2001.

[36] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed Energy Conservation for Ad-hoc Routing," in *the Proceedings of ACM/IEEE MOBICOM 2001*, 2001.

[37] V. Rodoplu and T. H. Ming, " Minimum Energy Mobile Wireless Networks," *IEEE Journal of Selected Areas in Communications*, vol. 17, no. 8, pp. 1333–1344, 1999.

[38] L. Li and J. Y. Halpern, "Minimum Energy Mobile Wireless Networks Revisited," in *the Proceedings of IEEE International Conference on Communications (ICC'01), Helsinki, Finland*, June 2001.

[39] J. H. Chang and L. Tassiulas, "Maximum Lifetime Routing in Wireless Sensor Networks," in *the Proceedings of Advanced Telecommunications and Information Distribution Research Program (ATIRP'2000), College Park, MD*, March 2000.

[40] K. Sohrabi, J. Gao, V. Allawadhi, , and G. Pottie, "Protocols for self organization of a wireless sensor network," *IEEE Pers.l Commun.*, vol. 7, no. 5, October 2000.

[41] K. Akkaya and M. Younis, "An Energy-Aware QoS Routing Protocol for Wireless Sensor Networks," in *the Proceedings of the IEEE Workshop on Mobile and Wireless Networks (MWN2003), Providence, Rhode Island*, May 2003.

[42] T. He, J. Stankovic, C. Lu, and T. Abdelzaher, "SPEED: A stateless protocol for real-time communication in sensor networks," in *the Proceedings of the International Conference on Distributed Computing Systems, Providence, RI*, May 2003.

[43] D. Johnson and D. Maltz, *Dynamic Source Routing in Ad Hoc Wireless Networks*. Kluwer Academic Publisher, 1996.

[44] C. Perkins, E. Belding-Royer, and S.Das, "Ad hoc on-demand distance vector (AODV) routing," in *the Proceedings of IEEE WMCSA'99*, February 1999.

[45] O. Chipara, Z. He, G. Xing, Q. Chen, X. Wang, C. Lu, J. Stankovic, and T. Abdelzaher, "Real-time Power-Aware Routing in Sensor Networks," in *the Proceedings of 14th IEEE International Workshop on Quality of Service (IWQoS 2006),New Haven, CT*, June 2006.

[46] P. Pothuri, V. Sarangan, and J. Thomas, "Delay-constrained, energy-efficient routing in wireless sensor networks through topology control," in *the Proceedings of 2nd IEEE International Conference On Networking, Sensing and Control*, April 2006.

[47] L. Yuan, W. Cheng, and X. Du, "An energy-efficient real-time routing protocol for sensor networks," *Computer Communications*, vol. 30, pp. 2274–2283, 2007.

[48] Z.Khalid, G. Ahmed, and N. Khan, "A Real-time Energy-aware Routing Strategy for Wireless Sensor Networks," in *the Proceedings of Asia-Pacific Conference on Communications*, 2007, pp. 381–384.

[49] S. Ergen and P. Varaiya, "Energy efficient routing with delay guarantee for sensor networks," *Wireless Networks, Springer, Netherlands*, June 16.

[50] M. Hamid, M. Alam, and H. C. Seon, "Design of a QoS-Aware Routing Mechanism for Wireless Multimedia Sensor Networks," in *the Proceedings of IEEE Global Telecommunications Conference*, 2008.

[51] A. Razzaque, M. Alam, M. Rashid, and C. S. Hong, "Multi constrained QoS geographic routing for heterogeneous traffic in sensor networks," in *the Proceedings of IEICE Transactions on Communications*, 2008, pp. 2589–2601.

[52] A. Ahmed and N. Fisal, "A real-time routing protocol with load distribution in wireless sensor networks," *Computer Commun.*, vol. 31, no. 14, pp. 3190–3203, 2008.

[53] J. Sen and A. Ukil, "An adaptable and QoS-aware routing protocol for Wireless Sensor Networks," in *the Proceedings of Wireless Communication,Vehicular Technology,Information Theory and Aerospace and Electronic Systems Technology*, 2009, pp. 767–771.

[54] A. Mahapatra, K. Anand, and D. P. Agrawal, "Qos and energy aware routing for real-time traffic in wireless sensor networks," *Computer Commun.*, vol. 29, pp. 437–445, February 2006.

[55] S. Ratnaraj, S. Jagannathan, and V. Rao, "OEDSR: Optimized energy-delay sub-network routing in wireless sensor network," in *the Proceedings of IEEE International Conference on Networking, Sensing and Control.*, 2006, pp. 330–335.

[56] M. Chen, V. Leung, S. Mao, and Y. Yuan, "Directional geographical routing for real-time video communications in wireless sensor networks," *Computer Communications*, vol. 30, pp. 3368–3383, 2007.

[57] I. Politis, M. Tsagkaropoulos, and T. Daguiklas, "Power efficient video multipath transmission over wireless multimedia sensor networks," *Mobile News*, vol. 13, pp. 274–284, April 2008.

[58] Y. Lan, W. Wenjing, and G. Fuxiang, "A Real-time and Energy Aware QoS Routing protocol for Multimedia Wireless Sensor Networks," in *the Proceedings of 7th IEEE International Conference On World Congress Control and Automation*, June 2008.

[59] Y. Sun, H. Ma, L. Liu, and Y. Zhang, "ASAR: An ant-based service-aware routing algorithm for multimedia sensor networks," *Front. Electr. Electron. Eng. China*, vol. 3, no. 1, pp. 25–33, 2008.

[60] P. Shanghong, S. Yang, S. Gregori, and T. Fengchun, "An adaptive QoS and energy-aware routing algorithm for wireless sensor networks," in *the Proceedings of 2008 International Conference on Information and Automation (ICIA)*, 2008, pp. 578–583.

[61] Z. Ke, L. Li, and N. Chen, "A Crossover Game Routing Protocol for Wireless Multimedia Sensor Networks," in *the Proceedings of Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008.*, August 2008, pp. 595–599.

[62] H. Haiping and W. Ruchuan, "Clustered-Control Algorithm for Wireless Multimedia Sensor Network Communications," in *International Conference on Communications and Mobile Computing*, April 2010, pp. 264–268.

[63] Y. M. Huang, M. Y. Hsieh, , and M. S. Wang, "Reliable transmission of multimedia streaming using a connection prediction scheme in clusterbased ad hoc networks," *Computer Communications*, vol. 30.

[64] Q. Lu, W. S. Luo, and Y. Zhang, "Two-Hop Clustered Image Transmission Scheme in Multimedia Sensor Networks," *Chinese Journal of Sensors and Actuators*, vol. 20.

[65] A. Cerpa and D. Estrin, "ASCENT: Adaptive Self-configuring Sensor Networks Topologies," *IEEE Trans. on Mobile Computing*, vol. 3.

[66] L. Shu, Y. Zhang, L. T. Yang, Y. Wang, and M. Hauswirth, "Geographic Routing in Wireless Multimedia Sensor Networks," in *the Proceedings of Second International Conference on Future Generation Communication and Networking*, 2008, pp. 68–73.

[67] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *the Proceedings of ACM/IEEE MOBICOM00, Boston, Massachusetts, USA*, 2000, pp. 243–254.

[68] S. Medjiah, T. T. Ahmed, and F. Krief, "GEAMS: A Geographic Energy-Aware Multipath Stream-based routing protocol for WMSNs," in *the Proceedings of Global Information Infrastructure Symposium, 2009*, June 2009, pp. 1–8.

[69] B. Hamdaoui and K. G. Shin, "OS-MAC: An efficient MAC protocol for spectrum-agile wireless networks," *IEEE Transactions on Mobile Computing*, August 2008.

[70] N. Chakchouk and B. Hamdaoui, "Traffic and interference aware scheduling for multi-radio multi-channel wireless mesh networks," *IEEE Transactions on Vehicular Technology*, to appear.

[71] C. Suh, Z. Mir, and Y. Ko, "Design and implementation of enhanced IEEE 802.15.4 for supporting multimedia service in Wireless Sensor Networks," *Special Issue on Wireless Multimedia Sensor Networks, Elsevier Computer Network Journal*, vol. 52, no. 13, pp. 2568–2581, September 2008.

[72] C. Li, P. Wang, H.-H. Chen, and M. Guizani, "A Cluster Based On-demand Multi-Channel MAC Protocol for Wireless Multimedia Sensor Networks," in *the Proceedings of IEEE International Conference on Communications*, 2008, pp. 2371–2376.

[73] O. D. Incel, S. M. P. Jansen, and P. Havinga, "MC-LMAC: A multi-channel MAC protocol for wireless sensor networks," *Ad Hoc Networks*, vol. 9, January 2011.

[74] N. Sabena, A. Roy, and J. Shin, "Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks," *Computer Networks*, vol. 52, no. 13, 2008.