International Journal of Innovative
Computing, Information and Control
Volume 8, Number 10(B), October 2012
 ISSN 1349-4198
pp. 7185–7198

A CROSS-LAYER PROTOCOL FOR RELIABLE AND EFFICIENT COMMUNICATION IN WIRELESS SENSOR NETWORKS

Weiwei Fang, Zhen Liu and Feng Liu
School of Computer and Information Technology
Beijing Jiaotong University
No. 3, Shangyuan Cun, Haidian District, Beijing 100044, P. R. China
{ wwfang; zhliu; fliu }@bjtu.edu.cn

Received September 2011; revised January 2012

Abstract. *It is a practical challenge to provide reliable and efficient communication in wireless sensor networks (WSNs). However, traditional layered protocols are not jointly designed and optimized to maximize the overall network performance while minimizing the energy expenditure. In this paper, we introduce a novel node initiative concept that allows the intrinsic protocol layer functionalities required for successful communication in WSNs to be implemented in a unified protocol framework. Based on this concept, a cross-layer, reliable and efficient communication protocol (CREC) is proposed, which implements a cross-layer operation of medium access contention, robust geographical routing, and distributed congestion control with due consideration on channel effects, information fidelity and energy efficiency to realize reliable and efficient data transmission in WSNs. Extensive simulation results show that CREC, though very simple, achieves significant improvements in terms of network performance and energy efficiency over other state-of-the-art solutions.*
Keywords: Wireless sensor network, Cross-layer design, Robust routing, Congestion control, Transmission reliability

1. **Introduction.** The technological advances in micro-electro-mechanical systems and wireless communications have motivated the development of wireless sensor networks (WSNs) [1]. A typical WSN is an event-driven system that exploits the collective efforts of densely deployed sensor nodes to continuously observe physical phenomenon and to reliably obtain sensory information. One of the key challenges in WSNs is how to achieve high transmission reliability under constraints of limited hardware resources (especially the energy constraints). Aside from careless design deficiencies, transmission reliability in WSNs is mainly influenced by multiple concurrent factors, such as channel error, link failure, network congestion, and energy exhaust. Meanwhile, it is required that reliability mechanisms should maintain data fidelity at an acceptable level [2]. Intuitively, it is a multifaceted problem that corresponds to the functionalities originally designed to be provided by different protocol layers, i.e., physical, medium access control (MAC), network, transport, and even application layers. However, the conventional layered protocols are individually developed and optimized for achieving high performance in terms of the metrics related to a certain networking layer. As a result, it is difficult or even impossible to simply and directly combine the protocols belong to different layers together to maximize overall network performance while minimizing node energy expenditure. Considering the constrained hardware resources of sensor nodes, cross-layer design [3], which exploits dependencies and interactions across layers to integrate multiple layers' functionalities into a unified communication framework, stands as the most promising technique to achieve reliable and efficient communication in WSNs [4].

A consistent number of recent works have focused on the cross-layer development and improvement of network protocols for WSNs [5, 6]. It has already been revealed that cross-layer design techniques result in significant improvements on network performance and energy efficiency in WSNs. However, as will be discussed in Section 2, these studies either provide analytical results without any practical protocol design or only focus on cross-layer interactions or modularity within a narrow scope, e.g., MAC and network layers. Apparently, they cannot fully fulfil reliability requirements for data transmission in WSNs. It is imperative to design a unified cross-layer protocol that considers upper layer functionalities with physical layer (wireless channel) effects for reliable and efficient communication in WSNs.

To this end, this paper introduces a novel "node initiative" concept, and illustrates how certain intrinsic functionalities required for successful communication in WSNs (i.e., medium access, robust routing and congestion control) can be unified based on this concept into a single protocol operation. Coupled with the receiver contention based relay mechanism, the node initiative concept provides a binary option for each node to decide on participation in data transmission according to its local current state related to reliability guarantees. Using the initiative concept, a novel protocol (CREC) is developed to implement a cross-layer operation of medium access contention, robust geographical routing, and distributed congestion control with consideration on physical channel effects, sensory data fidelity and energy efficiency optimization. CREC has been implemented along with several state-of-the-art cross-layer protocols in the ns-2 simulator. Extensive simulation results show that CREC significantly improves transmission performance as well as energy efficiency over existing works. These results highlight the advantages of the node initiative concept, which constitutes the core of CREC and represents a novel means for cross-layer design in WSNs.

The remainder of the paper is organized as follows. In Section 2, we briefly review existing works on cross-layer design in WSNs for reliability guarantees. Section 3 presents the CREC basics, overview, and protocol description in detail. The performance of the CREC protocol is evaluated via simulations in Section 4. Finally, the paper is concluded in Section 5.

2. Related Works.

In recent years, cross-layer protocol design has become a very prominent topic in network research and particularly in wireless area. In the following, we will provide a survey on some typical WSN protocols that provide reliable transmission services by cross-layer design. These studies are classified in terms of interactions or modularity among different protocol layers.

MAC + Physical: A multichannel MAC protocol Rainbow is described in [7], in which local time division multiple access (TDMA) and frequency hopping spread spectrum (FHSS) are jointly designed to control channel access. To meet the reliability needs for data collecting WSNs, the FHSS technique is used by Rainbow to evade radio frequency interference. However, many practical problems need to be solved before TDMA can be widely used in WSNs, including synchronization and scheduling overhead.

Routing + Physical: Traditional geographic routing protocols that employ a maximum-distance greedy forwarding technique perform poorly in realistic conditions with lossy links. Theses protocols have been improved by considering the channel quality information from the physical layer in routing decisions. For example, it is found that the product of the packet reception rate (PRR) and the distance improvement towards destination is a highly suitable metric for geographic forwarding in realistic environments [8]. However, the periodically PRR information exchange among sensor nodes my result in a high overhead in networks with high node density. Moreover, some important factors such as energy consumption and link asymmetry are not considered in [8].

Routing + MAC: In many works, the receiver-oriented routing is exploited for MAC and routing cross-layer modularity [9, 10]. Instead of pro-actively establishing the global end-to-end routing, this approach enables each neighbour of the current sender to dynamically contend for further relaying based on how well it is suited as the next-hop relay. Thus, the receiver-oriented routing is able to provide robust transmission routes against topology changes. However, most of these protocols employ very simple routing metrics and do not take link unreliability and network congestion into consideration.

Transport + Routing: There have been attempts to explore on-demand multi-path routing mechanisms for congestion alleviation in WSNs. Biased geographical routing (BGR) [11] is a geographic routing that reactively split traffic during congestion. However, the bias, which determines how far the trajectory of the splitting traffic will deviate from the original path, is randomly chosen and could make congestion worse in some cases. Congestion aware routing (CAR) [12] proposes to use a priority aware routing with data prioritization to alleviate congestion, requiring multiple sink nodes to be deployed for gathering data packets with different priorities.

Transport + Application: Some works focus on traffic regulation at application layer for congestion alleviation in WSNs. In event-to-sink reliable transport (ESRT) [13], the sink is able to regulate the source reporting rate in a uniformed way by broadcasting control message to all data sources. The underlying assumption is that a sink can reach all nodes via a high-energy one-hop broadcast, which is not practical for large-scale WSNs. Congestion control for sensor networks (CONCERT) [14] employs a adaptive data aggregation technique to pro-actively reduce the amount of data packets travelling throughout the network. However, it has been revealed that traffic reduction could impose a negative impact on data fidelity [15].

Multiple-Layer Solutions: In addition to those protocols that focus on pair-wise cross-layer design, some general cross-layer approaches among several protocol layers exist. A joint routing, MAC, and link layer optimization framework is proposed in [16]. However, the optimization problem is only theoretically solved but not further turned into a practical protocol implementation. Interference-minimized multipath routing (I2MR) [17] tries to avoid congestion by discovering multiple zone-disjoint routing paths with consideration on the effects of wireless interferences, and alleviate congestion by notifying source nodes to reduce loading rate. Channel-aware geographic-informed forwarding (CAGIF) [18] proposes a new local metric called efficient advancement metric (EAM) to solve the optimal relay selection problem in receiver-based geographic routing by adjusting the transmission ranges according to underlying channel conditions. CAGIF requires code division multiple access (CDMA) based wireless nodes which may not be suitable for WSNs, since the CDMA technology may not be the most efficient solution for WSN communications.

The studies above either provide analytical results without any protocol for practical implementation or perform cross-layer design within certain limited scopes. In this paper, we argue that a new networking paradigm is required to design a cross-layer protocol that addresses medium access, robust routing, and congestion control issues with consideration on channel effects, information fidelity and energy efficiency.

3. The CREC Protocol Design.

The core idea of CREC is to provide a unified framework where both the information and the functionalities of three fundamental communication paradigms, i.e., medium access, robust routing and congestion control, are jointly considered and implemented in a single protocol operation. The design details of CREC are presented in the following subsections.

3.1. **Models and assumptions.** Before presenting implementation specifics of CREC, we firstly introduce models and assumptions involved in this work.

A WSN composed of a large number of sensor nodes and a stationary sink s in an interested area is considered in this paper. Sensor nodes are randomly distributed in the area and will remain stationary after deployment. All of them have similar capabilities and equal significance. However, they are constrained in processing capability, memory space, energy storage, and communication bandwidth. Due to the limited radio range and energy constraints, multi-hop communication is exploited to relay sensory data from the source node to the sink. The sink has unlimited resources to perform data gathering and processing tasks. Motivated by the fact that WSN applications inherently require location information to uniquely identify a monitored object from all the other ones, we assume that each node is aware of its location via an on-board global positioning system (GPS) or a localization algorithm [19]. Moreover, the sink's location is pre-known to all sensor nodes from pre-programmed information, so that each of them can determine where to forward data packets. To efficiently utilize the shared wireless medium, all nodes in the network are assumed to coordinate medium access by adopting the carrier sense multiple access with collision avoidance (CSMA/CA) MAC protocol, which has been proven effective in many prior works [10, 13, 20]. The network model is geared toward event-based data flow, in which sensor nodes send data packet to the sink once a monitored object is detected in their vicinity. The size of data packet is assumed to be much larger than those of network control messages, e.g., Request-To-Send (RTS)/Clear-to-Send (CTS)/Acknowledgement (ACK) frames.

It must be noted that the idealized perfect-reception-within-range radio model is unlikely to be valid in any realistic environment. Therefore, we adopt a more realistic WSN radio model presented in [8]. In this model, the PRR in terms of the signal to noise ratio (SNR) γ is a random variable given by:

$$PRR(\gamma) = \left(1 - \frac{1}{2}\exp^{-\frac{\gamma B_N}{2R}}\right)^{\rho 8 f} \tag{1}$$

where B_N is the noise bandwidth, R is the data rate in bits, ρ is the encoding ratio, and f is the frame length. In this work, the node communication range, R_c, is defined as the maximum source-destination transmission distance within which $PRR > 0$.

3.2. **Node initiative for communication participation.** The node initiative concept coupled with the receiver-oriented packet forwarding mechanism provides freedom for each sensor node to choose whether to participate in communication or not. To provide a reliable and efficient end-to-end transmission service, the proposed concept should incorporate the intrinsic communication functionalities that are required to achieve successful communications in WSNs.

In CREC, the selection of the next-hop relay is performed by means of contention through RTS/CTS handshaking. For any node i which wants to initiate data transmission, it will firstly broadcast a RTS message to inform its neighbours that it currently has a packet to send. Upon receiving this message, each neighbour of node i decides whether to participate in the communication or not through the initiative assessment procedure. The initiative assessment is designed to be a binary operation in which a node decides to participate in communication only if its initiative value is equal to 1. The initiative, I, is determined based on the state of the sensor node's communication capabilities as follows:

$$I = H(l_{RTS} - l_{Th}) \cdot H(e_{Rem} - e_{Th}) \cdot H(\lambda_{Th} - \lambda_{Relay}) \cdot H(\beta_{Th} - \beta) \cdot H(\phi_{Th} - \phi) \tag{2}$$

where $H(x)$ is the Heaviside unit step function ($H = 0$ if $x < 0$ and $H = 1$ otherwise). Thus, the initiative I is set to 1 only if all the results of subtraction expression in Formula (2) are positive. Now we explain the meaning of each notation in these expressions in terms of its corresponding communication functionality:

- According to Formula (1), it is necessary to choose only those nodes with higher SNR to participate in communication since they can provide reliable links for data transmission. In WSNs, wireless transceivers commonly provide the Link Quality Indicator (LQI) reading [21] as a representation of the SNR value for practical implementation. Therefore, it is required that the LQI of a RTS message, l_{RTS}, is above some threshold l_{Th} for a node to participate in the communication.
- In order to extend the working lifetime of a sensor node, it is required that the node helps relay data packets only when its remaining energy, e_{Rem}, is above the minimum threshold, e_{Th}.
- To pro-actively avoid congestion, it is necessary to limit the input traffic that a sensor node is capable to relay. That is, the rate of overall relay traffic from neighbours, λ_{Relay}, must be below some threshold λ_{Th}.
- The buffer occupancy level of a node, β, should not exceed the maximum threshold, β_{Th}, so that the node will not suffer buffer overflow and the node-level congestion [22] can be prevented.
- To prevent the link-level congestion [22], a sensor node is triggered to periodically measure the channel loading when the buffer is not empty (i.e., $\beta > 0$). The channel loading, ϕ, should not exceed the maximum achievable channel utilization, ϕ_{Th}.

3.3. **Receiver contention for packet forwarding.** The receiver contention operation of CREC integrates the initiative assessment procedure within the receiver-oriented routing approach [9, 10]. When a node i has a data packet to transmit, it first broadcasts a RTS message, which contains the location of itself (As stated in Section 3.1, it is natural to leverage the node location information for communication). For any neighbour node that receives the RTS message from node i, it first performs initiative assessment as explained in Section 3.2. If a neighbour node j does not have the initiative to participate in the communication (i.e., $I = 0$), it switches to sleep state for the duration of current transmission to save energy. Otherwise, it is enabled to join the receiver contention for relaying the current packet from node i. Then node j generates a CTS message as a reply to the RTS message, and sets a proper delay, denoted by δ_{ij}, for broadcasting the CTS message based on a discrete delay function given in the next paragraph. If node j overhears a CTS message broadcasted by another candidate before δ_{ij} is due (Since the carrier sensing range is normally larger than twice of the transmission range, the CTS can be heard or sensed by all neighbours of node i), node j cancels broadcasting its own CTS message and switch to sleep; otherwise, node j broadcasts its CTS message when δ_{ij} is due. When node i receivers the CTS packet from a potential receiver j, it determines that the contention for packet relaying has ended and will unicast the packet to j. At the end, node j replies with an ACK message to node i after successfully receiving the packet.

It is possible that several nodes from neighbours of node i can become relay candidates, especially when the network density is relatively high. To reduce potential collisions as well as communication overhead incurred by relay selection, a discrete delay function is designed in CREC to promote the best relay and suppress the broadcasting of CTS message by other candidate neighbours. While many policies can be used to decide this function according to the application specifics, without loss of generality, we consider the following metrics related to the node state in our CREC protocol:

- Node-to-sink distance: The next-hop relay should be selected from nodes that are closer to the sink with respect to the current node, i.e., $d_{js} < d_{is}$ (d denotes the Euclidean distance).
- Residual node-energy: To balance energy consumption, the node carrying more residual energy should be given priority in the contention.
- Distance-hop trade-off: There exists contradiction between forwarding distance and energy consumption in geographic routing protocols [8]. If the protocol tries to minimize the number of overall hops by maximizing the geographic distance covered at each single hop, it is likely to incur considerable energy expenditure due to retransmission on the unreliable long weak links. On the other hand, if the protocol tries to maximize per-hop reliability by forwarding only to close neighbours with good links, it may cover only a small geographic distance at each hop, which would also result in significant energy expenditure due to the need for more transmission hops for the data packet to reach the sink. It has been suggested in [8] that the production of PRR and the sender-receiver distance is an optimal metric for balancing distance-hop trade-off in lossy wireless networks.

To incorporate the metrics discussed above, we define a weighted additive function to set the back-off delay for a neighbouring node j of node i as follows:

$$\delta_{ij} = \frac{\left[w_1 \cdot \left(1 - \frac{d_{is} - d_{js}}{R_c}\right) + w_2 \cdot \left(1 - \frac{e_j}{E}\right) + w_3 \cdot \left(1 - \frac{PRR(i,j) \cdot d_{ij}}{R_c}\right) + w_4 \cdot v\right] \cdot SIFS}{\sum_{k=1}^{4} w_k} \tag{3}$$

where e_j and E denotes the residual and initial energy of node j respectively, v is a random value between 0 and 1, $SIFS$ is the standard Short Inter-Fame Spacing in CSMA MAC, and w_k ($k \in [1, 4]$) are the weighting coefficients used to weight among application requirements on the metrics discussed above while w_4 is set as the smallest one of the four coefficients to resolve the CTS collisions among several contenders. It is obvious that in the receiver contention the neighbour node with the shortest back off time for CTS reply will be the final next-hop relay for the data packet from node i.

Note that node i will not receive any CTS packet if there is no qualified relay candidate that has a shorter distance to sink and an initiative to communicate (i.e., $I = 1$). If no response is received after t (e.g., 3 times) retries, node i determines that a local minimum is encountered, and switches to angle-based routing as explained in the next sub-section.

3.4. **Angular relaying for void traversal.** The communication void problem arises when a node cannot find any feasible neighbour that is closer to the sink than itself and has the initiative to participate in the communication. One possible solution is to deploy as many sensor nodes as possible [23]. However, this approach may be impractical in many scenarios since the cost is too high to afford. Face routing [24] is another technique proposed for resolving the void problem. It requires information exchange among the neighbours of the node to establish a planarized graph and construct routes to traverse around the void, thus inducing heavy communication overhead for energy-constrained sensor nodes [25]. Therefore, we introduce a stateless angle-based routing technique to address the communication void problem.

We use a void topology in Figure 1 to illustrate the main principle of angular relaying in CREC. When a data packet reaches a local minimum at node i, it has to be routed around the void either in clockwise direction (though node such as a or c) or in counter-clockwise direction (through node such as b or d). To minimize the total hops of a detour path, it is better to route the data packet along the perimeter of the void through nodes in closest proximity to the void, e.g., node a (clockwise direction) or b (counter-clockwise direction).

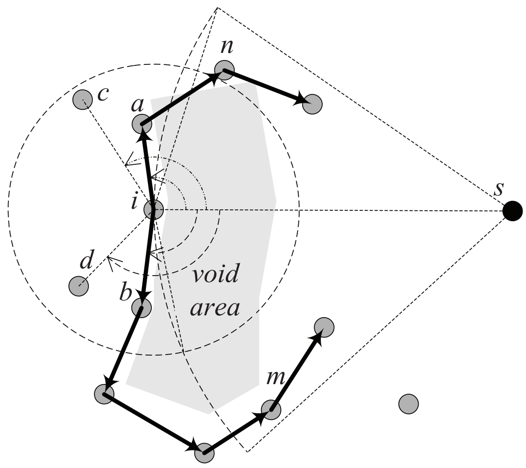

FIGURE 1. Illustration of angle-based routing for void traversal in CREC

It can be noticed from Figure 1 that a perimeter neighbour has the smallest deflection angle (denoted by θ) between the detour direction and the line is among all neighbours of node i, e.g., $\theta_{sia} < \theta_{sic}$ (clockwise direction) and $\theta_{sib} < \theta_{sid}$ (counter-clockwise direction). Based on this geometric property, we can design a new angle-based delay function for receiver contention in the angle-based routing mode.

To balance traffic load for perimeter nodes around the void, the source node is responsible for randomly selecting a potential traversal direction (clockwise or counter-clockwise) for each data packet it generates. This direction selection is carried by the data packet and used by all the local minimum nodes (e.g., node i in Figure 1) for void traversal. When node i switches to angle-based routing mode, it broadcasts a special RTS message, which uses certain bits in the reserved frame space to indicate both the routing mode and the selected traversal direction. Upon receiving this RTS message from node i, each neighbour node with the initiative (i.e., $I = 1$) sets the back-off delay for the CTS message replying according to an angle-based delay function. Specifically, the delay function for a neighbouring node j is given as follows:

$$\delta'_{ij} = \frac{\left[w_5 \cdot \left(\frac{\theta_{sij}}{2\pi}\right) + w_6 \cdot \left(1 - \frac{e_j}{E}\right) + w_7 \cdot \left(1 - \frac{PRR(i,j) \cdot d_{ij}}{R_c}\right) + w_8 \cdot v\right] \cdot SIFS}{\sum_{k=5}^{8} w_k} \tag{4}$$

where w_5 is set as the largest one of the four weighting coefficients and the other w_k ($k \in [6, 8]$) are used to distinguish several contenders with equal angles, and $\theta_{sij} \in (0, 2\pi)$ can be calculated through the cosine law with coordinate transformation techniques [26]. Apparently, the neighbouring node firstly replying with the CTS message will be the next-hop relay for the current data packet forwarded by node i. The procedure above is repeated until the data packet finally reaches a node that is closer to the sink (e.g., node m or n) than the node that initiated angular relaying (e.g., node i). Then the void traversal is ended, and CREC switches back to the basic routing mode.

However, the void traversal mechanism introduced above cannot prevent routing loops that are created in the case when all candidate neighbours are located at the opposite side to the detour direction (i.e., $\theta_{sij} \in (\pi, 2\pi)$). As shown in Figure 2, both θ_{sia} and θ_{sic} are greater than π because the only two neighbours of node i, i.e., node a and c, are not located in the selected counter-clockwise direction. Node a is chosen as the next-hop relay for the data packet from node i because $\theta_{sia} < \theta_{sic}$. However, the data packet will be routed from node back to node i via node c since $\theta_{sac} < \theta_{sai} < \theta_{sab}$ and $\theta_{sci} < \theta_{sca}$.

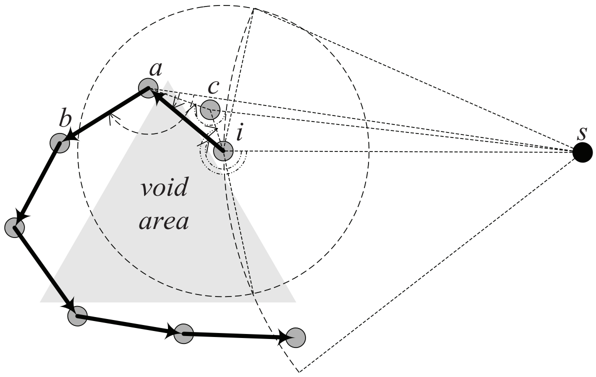

FIGURE 2. A void traversal scenario in which a routing loop may appear at node i

Suppose node c does not exist, the data packet will be directly routed from node a back to node i since $\theta_{sai} < \theta_{sab}$. Thus, it is necessary for the protocol to prevent nodes located on the clockwise side of line ia (i.e., node a and c) to respond to a special RTS message from node a. To achieve this, CREC provides a complementary policy for loop avoidance: if a sensor node j becomes the next-hop relay for a data packet from node i while $\theta_{sij} > \pi$, then the special RTS message from node j for this data packet should contain the angle θ_{sji} to prevent any neighbour of node j (e.g., node k) with a deflection angle (i.e., θ_{sjk}) not greater than θ_{sji} from replying to this RTS message.

3.5. **Traffic regulation for congestion control.** A sensor node has two duties in WSNs, i.e., data source and packet router. Therefore, WSNs exhibit a unique funnelling effect where sensory data generated in the sensing area travel hop-by-hop in a many-to-one traffic pattern towards the sink node, resulting in traffic intensity and sometimes even congestion as a surge of data packets move closer to the sink. Congestion has dreadful consequences in terms of packet loss, energy efficiency and fidelity degradation in WSNs. With the help of node initiative concept, CREC incorporates distributed traffic control mechanisms to resolve the congestion for both the packet router and the data source duties of a sensor node.

We first analyse the upper bound for node relaying load, λ_{Th}, which is used for congestion avoidance in the CREC initiative assessment as presented in Formula (2). For any node i, its overall input and output packet rate can be represented respectively as:

$$\lambda_{i,In} = \lambda_{i,Src} + \lambda_{i,Relay} \tag{5}$$

$$\lambda_{i,Out} = (1 + \mu_i)\lambda_{i,In} \tag{6}$$

where μ_i is the packet error rate that is calculated as a moving average of the packet loss rate encountered by node i, and $(1 + \mu_i)$ is used to approximate the actual number of transmission times for all the packets in the buffer. Note that $\lambda_{i,Out}$ is higher than $\lambda_{i,In}$ because node i attempt to retransmit the packet that were not successfully sent to the next hop. According to (5) and (6), the average time duration spent by node i for receiving and transmitting in a considerable long interval T are given respectively as:

$$T_{tx} = T \cdot \lambda_{i,Out} \cdot T_{Packet} \tag{7}$$

$$T_{rx} = T \cdot \lambda_{i,Relay} \cdot T_{Packet} \tag{8}$$

where T_{Packet} is the average duration for node i to transmit a packet to the next hop.

To prevent congestion at node i, all the packets that generated or relayed by it should be transmitted in interval T. Therefore, it is required that

$$T \geq T_{tx} + T_{rx} \tag{9}$$

Combining Formulas (5)-(9), we finally have

$$\lambda_{i,Relay} \leq \frac{1}{2+\mu_i} \cdot \left[\frac{1}{T_{packet}} - (1+\mu_i) \cdot \lambda_{i,Src} \right] = \lambda_{i,Th} \tag{10}$$

According to Formula (2), node i participates in packet relaying as long as Formula (10) is satisfied. It can be noticed that the nodes at the traffic hotspot (i.e., nodes with relatively longer transmission duration T_{packet}) has a relatively lower threshold for packet relay rate. Formula (10) also ensures that the packet relay rate of source nodes (i.e., nodes with $\lambda_{Src} > 0$) is lower than that of the nodes which act as only relays (i.e., $\lambda_{Src} = 0$). Both of these help to maintain homogeneous distribution of traffic load throughout the network. Note that, node i timely updates $\lambda_{i,Th}$ with input parameters μ_i, T_{packet} and $\lambda_{i,Src}$ after a successful or unsuccessful transmission of a data packet.

Formula (10) attempts to provide evenly distribution of traffic load in the network so as to prevent the potential congestion in the long term. In some cases, traffic hotspots may still form due to some short-term traffic changes which are difficult to be predicted by individual sensor nodes. According to Formula (2), if a node detects the onset of congestion by β or ϕ, it stops to participate in the communication immediately.

In addition to regulating packet relay as discussed above, CREC also takes a local congestion control operation by directly regulating the amount of source traffic injected into the network. That is, if node i cannot receiver any CTS reply in the receiver contention while it can still overhear the RTS message from some neighbouring nodes, it infers that itself is located in the traffic hotspot. In this case, node i directly reduces the reporting rate of generated packets $\lambda_{i,Src}$ to decrease the traffic load [26]. Generally, a source node follows a regulation policy that is similar to the popular Additive Increase Multiplicative Decrease (AIMD) mechanism in Transmission Control Protocol (TCP) [22]:

$$\lambda_{Src}(t+\Delta t) = \begin{cases} \lambda_{Src}(t) + \Delta\lambda, & \text{if } \lambda_{Src} < \lambda_{Def} \\ \lambda_{Def}, & \text{if } \lambda_{Src} = \lambda_{Def} \\ \lambda_{Min}, & \text{if congestion is detected} \end{cases} \tag{11}$$

where $\Delta\lambda$ and Δt are the constants for rate increment and update interval respectively, λ_{Def} and λ_{Min} are the default reporting rate and the tolerable minimal reporting rate respectively. All these values are artificially pre-configured for all the sensor nodes before network deployment on the basis of data characteristics and application specifics [15, 22]. It must be noted that such a control mechanism is applied by some node located in the traffic hotspot region only to its own generated packet rate when all its neighbours refused to relay packets according to the initiative assessment in Formula (2).

The reduction of source traffic, though very effective for congestion alleviation, may probably degrade application fidelity at the sink if some key sampling values are neglected without reporting. Thus, the source node needs to select representative sampling data but not random or periodic selected sampling data as event reports sent to the sink [22]. In this way, information fidelity can be well maintained at the sink while source traffic is regulated for congestion alleviation. To this end, CREC can integrate application-layer data processing techniques to perform selective reporting based on data characteristics. Take numeric measures such as temperature for example, the source node can use the piecewise linear approximation technique to approximate the time series with a sequence of line segments [15]. More investigations about relevant techniques are out of the scope of this paper.

4. **Performance Evaluation.** To gain more insight into the protocol operation and performance, we have implemented a simulation package based on ns-2. We compare CREC against another two existing cross-layer protocols: SGF [10] and SPEED [23]. SGF is a state-free receiver-based forwarding protocol that also employs the MAC layer contention for next-hop relay selection. It adopts the gradient routing mechanism to avoid communication void. However, the transport layer issues such as congestion control are not taken into account by SGF. SPEED uses the relay speed metric for per-hop relay selection in geographical forwarding to achieve soft real-time delivery and traffic load balance. It employs a back-pressure beaconing mechanism to handle both void and congestion problems. Considering these design features, we choose to compare our work against these two protocols to evaluate the performance of CREC. Note that SGF in [10] did not take channel condition into account, so we revised the delay timer setting in SGF by adding PRR as one of the routing metrics. Besides, hop count is selected as gradient cost since node transmitting power is non-adjustable in this work.

Refer to [8, 21, 22, 26] for parameters close to those of a real sensor node, the simulation settings and specific parameter values are given as in Table 1.

Table 1. Simulation settings and parameters

Parameter	**Value**	**Parameter**	**Value**
Sensing Area Size	500m × 500m	l_{Th}	$65 \in [50, 100]$
Sink Location	$(500, 500)$	e_{Th}	100μJ
Buffer Length	30	β_{Th}	70%
Transmission Bandwidth	250kbps	ϕ_{Th}	70%
Traffic Type	Constant Bit Rate (CBR)	$f_{control}$	20bytes
Transmitting Energy	24.75mW	f_{data}	250bytes
Receiving Energy	13.5mW	R_c	30m
Sleeping Energy	15μW	$SIFS$	10μs

In the first set of experiments, we compare and evaluate the transmission performance of the protocols in light-loaded network scenarios. All sensor nodes are evenly distributed in the whole area, and data flows are randomly generated from some source nodes. Figure 3 and Figure 4, respectively, show the packet energy consumption per hop and the average transmission delay per hop with respect to different network densities. We could find out that the performance differences between SGF and CREC are quite slight in the low-density networks. However, SGF costs additional energy and time to establish network gradient and update it when necessary. In CREC, the node location is exchanged among neighbouring nodes only when the sender initiates the data forwarding through RTS message broadcast. Due to these reasons, we can notice that the energy difference between them becomes relatively larger as the node density grows higher. Meanwhile, we can find that SPEED performs the worst among the three due to the inefficiency of its packet relaying mechanism. Moreover, the periodic beaconing of node state information in SPEED incurs considerable overhead of energy and delay. Additionally, the results imply that it is necessary to maintain a moderate network density to select a more optimal forwarder to guarantee the real-time data transmission in CREC, although more energy overhead may be incurred for that.

In the second set of experiments, we randomly create communication voids in the low-density networks (i.e., totally 500 nodes) by manually removing some deployed nodes

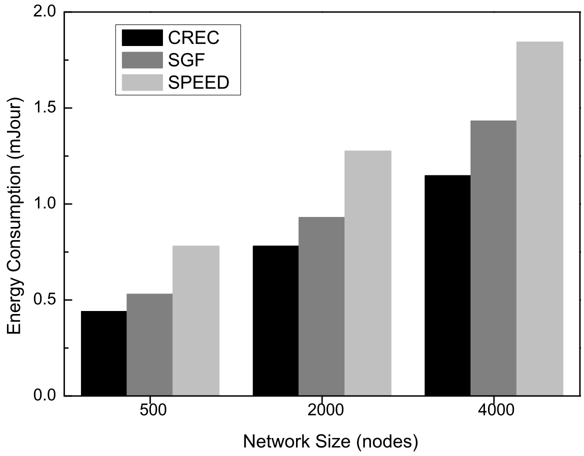

FIGURE 3. Packet energy consumption per hop versus different node densities

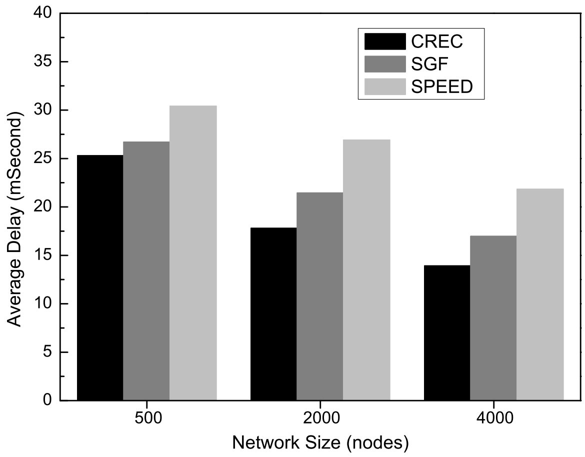

FIGURE 4. Average transmission delay per hop versus different node densities

from them. Since SGF is free of void problem by setting up a tree-like transmission architecture based on the cost gradient, we choose SGF as a baseline for the comparison on void handling capability. Figure 5 shows the comparison results on path length in void scenarios, in which each result has been normalized to that of SGF. From Figure 5, CREC performs much better than SPEED, but a little inferior to SGF in terms of hop count. The reason is that in CREC the packet tries to traverse along the border of communication void through the angular routing. Unlike SGF, CREC does not have information such as gradient cost to minimize the length of a detour path. However, we argue that the overhead for maintaining the cost filed is significantly high, especially in the large-scale networks. On the other hand, SPEED handles the void problem through a backtracking like forwarder searching method, which considerably increases the total bypassing hops for void traversal.

In the third set of experiments, we compare and evaluate the transmission performance of the protocols in over-loaded network scenarios. To achieve this, we take a similar method to the one in [22] for creating traffic hot spots in a 2000-node network by simultaneously initiate four data flows from sources to a sink. Figure 6 and Figure 7, respectively, show the end-to-end delivery ratio and the energy consumption per packet with respect to

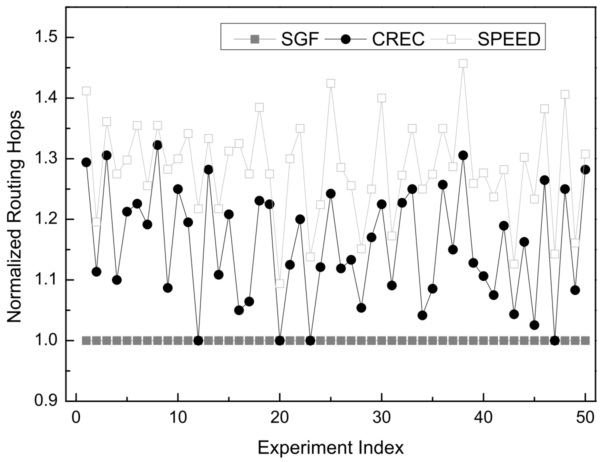

FIGURE 5. Comparison on routing path length in the void scenarios

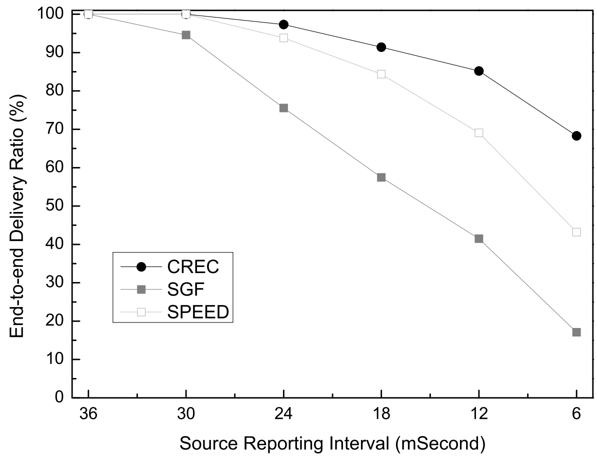

FIGURE 6. End-to-end delivery ratio versus different source reporting rates

different reporting rates. Since SGF is designed without any congestion handling mechanism, we can observe that it suffers significant data losses due to traffic overload. On the contrary, both SPEED and CREC have their own techniques for load balancing and congestion alleviation. Thus, we can observe that the energy consumption of SGF increases sharply and quickly exceeds that of SPEED when traffic load keeps on increasing. Since SPEED does not regulate source behaviour, it has relatively higher loss rate and energy consumption than CREC in the case of heavy congestion. Though in CREC the total network throughput would be reduced to some extent by source regulation, we argue that it is worth while performing congestion control to achieve energy efficiency in WSNs. As stated in Section 3.5, the selective reporting mechanism can help CREC to reduce the effect of fidelity degradation and present as accurate as possible data report to the sink.

5. **Conclusions.** In this paper, we have introduced a novel node initiative concept that enables multiple communication functionalities to be implemented in a unified protocol framework. Based on this concept, the cross-layer protocol CREC is proposed to provide the networking functionalities of medium access contention, robust geographical routing, and distributed congestion control for WSNs, with consideration on physical channel effects, sensory data fidelity and energy efficiency optimization. Simulation results show

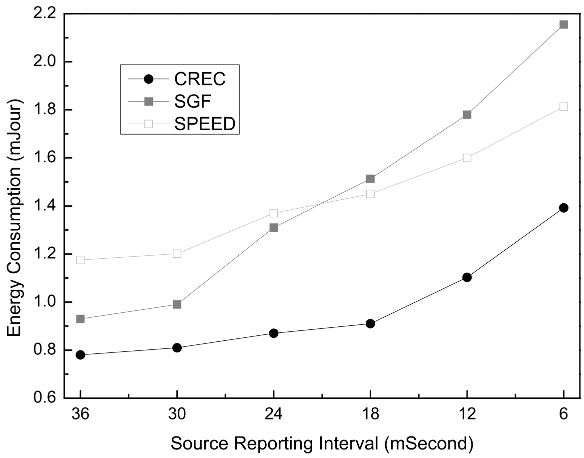

Figure 7. Average energy consumption per packet versus different source reporting rates

that the proposed protocol significantly improves the network performance, and outperforms several previous cross-layer protocols (i.e., SGF and SPEED) in the considered network scenarios. To the best of our knowledge, this is the first work that integrates functionalities of all layers, from physical to application, into a unified cross-layer protocol to achieve reliable and efficient communication in WSNs.

In the future, we plan to take effects induced by some control functionalities, such as sleep scheduling [20, 23] and topology control [1], into consideration to further extend and improve the current work. In addition, we will implement and evaluate CREC on a real sensor network test-bed.

Acknowledgement. This work is partially supported by the China Postdoctoral Science Foundation under Grant No. 20110490282, and the Fundamental Research Funds for the Central Universities of China under Grant No. 2011JBM225. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, A survey on sensor network, *IEEE Communications Magazine*, vol.40, no.8, pp.102-144, 2002.
[2] Y. F. Zhou, R. M. Lyu, J. C. Liu and H. Wang, PORT: A price-oriented reliable transport protocol for wireless sensor networks, *Proc. of the 16th IEEE Symp. on Software Reliability Engineering*, Chicago, IL, USA, pp.117-126, 2005.
[3] V. Srivastava and M. Motani, Cross-layer design: A survey and the road ahead, *IEEE Communications Magazine*, vol.43, no.12, pp.112-119, 2005.
[4] T. Melodia, M. C. Vuran and D. Pompili, The state of the art in cross-layer design for wireless sensor networks, *Proc. of 2005 EuroNGI Wksp. on Wireless and Mobility*, Como, Italy, pp.78-92, 2005.
[5] N. P. Mahalik, *Sensor Network and Configuration: Fundamentals, Techniques, Platforms, and Experiments*, Springer-Verlag, Berlin, Germany, 2007.
[6] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*, John Wiley & Sons Ltd., Sussex, Britain, 2010.
[7] Y. Yang and W. D. Yi, Rainbow: Reliable data collecting mac protocol for wireless sensor networks, *Proc. of 2010 IEEE WCNC Conf.*, Sydney, Australia, pp.1-6, 2010.
[8] M. Z. Zamalloa, K. Seada, B. Krishnamachari and A. Helmy, Efficient geographic routing over lossy links in wireless sensor networks, *ACM Transactions on Sensor Networks*, vol.4, no.3, pp.1-33, 2008.
[9] J. Zorzi and R. Rao, Geographic random forwarding (GeRaf) for ad hoc and sensor networks: Multihop performance, *IEEE Transactions on Mobile Computing*, vol.2, no.4, pp.337-348, 2003.

[10] P. Huang, H. Y. Chen, G. L. Xing and Y. D. Tan, SGF: A state-free gradient-based forwarding protocol for wireless sensor networks, *ACM Transactions on Sensor Networks*, vol.5, no.2, pp.1-25, 2009.

[11] L. Popa, C. Raiciu, I. Stoica and D. Rosenblum, Reducing congestion effects in wireless networks by wireless networks by multipath routing, *Proc. of the 14th IEEE ICNP Conf.*, Santa Barbara, CA, USA, pp.96-105, 2006.

[12] R. Kumar, R. Crepaldi, H. Rowaihy, A. F. Harris, G. H. Cao, M. Zorzi and L. T. Porta, Mitigating performance degradation in congested sensor networks, *IEEE Transactions on Mobile Computing*, vol.6, no.6, pp.682-697, 2008.

[13] O. B. Akan and I. F. Akyildiz, Event-to-sink reliable transport in wireless sensor networks, *IEEE Transactions on Networking*, vol.13, no.5, pp.1003-1016, 2005.

[14] L. Galluccio, A. Campbell and S. Palazzo, CONCERT: Aggregation-based congestion control for sensor networks, *Proc. of the 3rd ACM SENSYS Conf.*, San Diego, CA, USA, pp.274-275, 2002.

[15] C. Liu, K. Wu and J. Pei, An energy-efficient data collection framework for wireless sensor networks by exploiting spatiotemporal correlation, *IEEE Transactions on Parallel and Distributed Systems*, vol.18, no.7, pp.1010-1023, 2007.

[16] S. G. Cui, R. Madan, A. Goldsmith and S. Lall, Joint routing, mac, and link layer optimization in sensor networks with energy constraints, *Proc. of 2005 IEEE ICC Conf.*, Seoul, South Korea, pp.725-729, 2005.

[17] J. Y. Teo, Y. J. Ha and C. K. Tham, Interference-minimized multipath routing with congestion control in wireless sensor network for high-rate streaming, *IEEE Transactions on Mobile Computing*, vol.7, no.9, pp.1124-1137, 2008.

[18] L. L. Zhang and Y. Zhang, Energy-efficient cross-layer protocol of channel-aware geographic-informed forwarding in wireless sensor networks, *IEEE Transactions on Vehicular Technology*, vol.58, no.6, pp.3041-3052, 2009.

[19] J. Wang, R. K. Ghosh and S. K. Das, A survey on sensor localization, *Journal of Control Theory and Applications*, vol.8, no.1, pp.2-11, 2010.

[20] W. Ye, J. Heidemann and D. Estrin, An energy-efficient mac protocol for wireless sensor networks, *Proc. of the 21th IEEE INFOCOM Conf.*, New York, NY, USA, pp.1567-1576, 2002.

[21] G. Carles, B. Antoni and P. Josep, Impact of LQI-based routing metrics on the performance of a one-to-one routing protocol for IEEE 802.15.4 multihop networks, *EURASIP Journal on Wireless Communications and Networking*, no.205407, pp.1-20, 2010.

[22] W. W. Fang, J. M. Chen, L. Shu, T. S. Chu and D. P. Qian, Congestion avoidance, detection and alleviation in wireless sensor networks, *Journal of Zhejiang University Science C*, vol.11, no.1, pp.63-73, 2010.

[23] T. He, J. A. Stankovic, T. F. Abdelzaher and C. Lu, A spatiotemporal communication protocol for wireless sensor network, *IEEE Transactions on Parallel and Distributed Systems*, vol.16, no.10, pp.995-1006, 2005.

[24] D. Z. Chen and P. K. Varshney, A survey of void handling techniques for geographic routing in wireless networks, *IEEE Communications Surveys and Tutorials*, vol.9, no.1, pp.50-67, 2007.

[25] M. C. Vuran and I. F. Akyildiz, XLP: A cross-layer protocol for efficient communication in wireless sensor networks, *IEEE Transactions on Mobile Computing*, vol.9, no.11, pp.1578-1591, 2010.

[26] W. W. Fang, X. Zhao, T. S. Chu, Y. An, Y. Liu and D. P. Qian, RRDD: Receiver-oriented robust data delivery in mobile sensor networks, *Proc. of the 6th IEEE MASS Conf.*, Macau SAR, China, pp.876-881, 2009.