

A Distributed Cache Points Selection Scheme for Reliable Transport Protocols with Intermediate Caching in Wireless Sensor Networks

Mehdi Zirak*, Mohammad Hossein Yaghmaee*, Seyed Reza Kamel Tabbakh*

* Computer Engineering Department, Islamic Azad University, Mashhad, Iran
mehdi.zirack@gmail.com, yaghmaee@ieee.org, drkamel@mshdiau.ac.ir

Abstract— Traditional reliable transport protocols use end-to-end retransmission for the recovery of lost packets, which has high cost in Wireless Sensor Networks (WSNs). Hence, reliable transport protocols in WSNs use caching packets at the intermediate nodes in order to avoid costly end-to-end retransmission for loss recovery. Hop-by-hop retransmission scheme is one of the most important caching policies. This scheme requires large cache memory space which is in paradox with memory constraint in sensor nodes. In this paper, we propose a Distributed Cache Points Selection (DCPS) scheme. The proposed scheme tries to select cache points using a model based on probability theory so that it prevents packets end-to-end retransmission. Evaluation results show the efficiency of the proposed scheme regarding the required cache memory space.

Keywords— Wireless sensor networks, Reliable transport protocols, Loss recovery, Intermediate caching, Cache points selection.

I. INTRODUCTION AND RELATED WORKS

Sensor nodes are small, inexpensive, and low power devices with the capability of wireless communications. These nodes are able to perform some tasks such as sensing physical information (e.g. temperature, light, and sound) and data processing independently. A Wireless Sensor Network (WSN) is a collection of independent sensor nodes, which collaborate together distributedly, for the sake of monitoring a physical environment and gathering the environmental information. Packet loss is inevitable in WSNs because of bit error, congestion, and interference, whereas many of applications in WSNs are sensitive to the packet loss. Consequently, reliability guarantee in loss sensitive applications has become a challenging problem in WSNs.

Generally, transport layer is responsible for reliable packet delivery [1]. Reliable transport protocols for reliability guarantee must detect loss packets on intermediate links and recover them. Traditional reliable transport protocols (e.g. TCP) for loss recovery use end-to-end retransmission. In this scheme, only the source is responsible for the detection and recovery of lost packets [2]. Unlike wired networks, using end-to-end retransmission scheme in WSNs has low efficiency. This is due to lossy link environmental and the nature of multihop transmission in WSNs. Because, based on the Probability theory, the error rate is aggregated on several

hops in an exponential form. This condition causes the final destination located in the multihop distance from the source to have lower chance of intact receiving of packets sent by the source.

In order to avoid end-to-end retransmission, transport protocols were proposed which use intermediate caching where intermediate nodes are able to cache packets and retransmit them on-demand [3]. Intermediate caching leads to the packets retransmission point moves to a node that is closer to the sink rather than the source. In the protocols benefiting from intermediate caching, intermediate nodes don't consider priority for caching among the packets received from different flows. Some of these protocols consider memory constraint of intermediate nodes such as [4] and [5]. Whereas, most of them don't consider memory constraint and use hop-by-hop retransmission scheme. In the hop-by-hop scheme, each packet should be cached by each of the intermediate nodes [6]. This scheme, especially in the intermediate nodes where many concurrent flows pass through (e.g. the nodes near the sink), will require huge cache memory space. The [2], [7], [8], and [9] are the most well-known transport protocols that use hop-by-hop scheme. Given the limited memory in sensor nodes, packets may only need to be cached at selected nodes [6].

In this paper, we propose a Distributed Cache Points Selection (DCPS) scheme where the packets of a flow are cached at the selected intermediate nodes as cache points. In the proposed scheme, given the condition of flows that pass through an intermediate node, it only caches the packets of special flows. DCPS scheme is based on the probability theory. The proposed scheme tries to improve the efficiency of intermediate caching scheme considering the priority of flows for caching and also the memory constraint of intermediate nodes.

The rest of this paper is organized as follows. Section 2 describes the DCPS scheme. Section 3 presents the simulation results. The paper conclusion and future work are presented in Section 4.

II. PROPOSED SCHEME

The flows passing through an intermediate node often do not experience the same conditions. For example the ones having the most distance from the sink and higher error rate in

their intermediate links have more difficult conditions for delivering their packets to the sink. As mentioned in part 1, reliable transport protocols use intermediate caching to avoid costly end-to-end retransmission. However, considering the memory constraint in sensor nodes, an intermediate node will not be able to cache all received packets from different flows.

The flows with different conditions and memory constraints in the intermediate nodes requires a scheme in which each intermediate node will be selected as cache point only for limited flows. Cache point selection must be so that the greatest decrease happens in end-to-end retransmissions. Hence, before introducing the proposed DCPS scheme, we must look for the answer to this problem, in order to adopt an appropriate scheme to determine the cache points using this parameter: which parameter has the most impact on increasing the number of end-to-end retransmission for the delivery of a packet from source to sink?

A. The Effective Parameter

The nature of data transmission in WSNs is hop-by-hop. Hence, based on the probability theory, if a node sends a packet for destination in the n hop farther, the probability of receiving the message by destination is presented in Eq. 1. A summary of the mathematical notations is provided in Table 1.

$$PoR_n = \prod_{k=0}^{n-1} (1 - PER_{k,k+1}) \quad (1)$$

TABLE 1. SUMMARY OF MATHEMATICAL NOTATIONS

Notation	Meaning
$PER_{i,j}$	The Packet Error Rate on link between the nodes i and j .
PoR_n	Probability of Receiving of a packet by destination in the distance of n hop.
F_i^n	Total number of flows that cross node i .
C_n	The number of flows whose packets are cached in an intermediate node.
φ_i^f	Probability of receiving packets of flow f from the previous source at node i .
L_f	The number of end-to-end hop of flow f .
T_φ	The threshold of probability of receive.

In the end-to-end retransmission scheme, the source must retransmit the lost packets in the path. Hence, the probability of being received at least once by destination, for a packet after t times being sent by source, is calculated as follow:

$$r = \sum_{k=0}^{t-1} (1 - PoR_n)^k \cdot PoR_n \quad (2)$$

In this equation, $(1 - PoR_n)^k$ is the probability of not receiving a packet after k times being sent by source and PoR_n is the probability of successful receiving of the packet in $k+1$ sending. Eq. 2 shows, after t times sending a packet by source, the packet is expected to be received with minimum reliability r in the destination with distance of n hops. Lemma (1) shows the proposed geometric series in Eq. 2 in the form of an equation with degree t .

Lemma (1):

$$r = \sum_{k=0}^{t-1} (1 - PoR_n)^k \cdot PoR_n = \frac{(1 - (1 - PoR_n)^t) \cdot PoR_n}{1 - (1 - PoR_n)} = 1 - (1 - PoR_n)^t \quad (3)$$

Eq. 4, using Eq. 3 calculates the maximum number of the required transmissions (including retransmission) to achieve the end-to-end reliability r .

$$t = \log_{(1 - PoR_n)}(1 - r) \quad (4)$$

For cache points selection so that leads to the reduction of the number of end-to-end retransmission, we must look for a parameter in Eq. 4 having the maximum impact on the number of end-to-end transmission (t). In Eq. 4, the number of required end-to-end transmissions to deliver a packet to the sink grows rapidly, with the reduction of PoR for a constant value of r . So, we can conclude that intermediate nodes must try to cache packets of flows with less PoR in order to reduce the end-to-end retransmission.

B. Distributed Cache Points Selection

In this section with the aid of the impact parameter introduced in section 3.1 namely probability of receive (PoR), we propose a distributed cache points selection scheme for flows of a WSN. In the proposed scheme, an intermediate node is selected as cache point for flows having lower PoR among the flows passing through it. We have assumed that All the packets of a flow cached in a cache point can be recovered by cache point if they are lost during the path. Hence, we can consider a cache point as a new resource for cached packets from a flow.

To explain DCPS scheme let $\varphi_i^f, f \in \{1, \dots, F_i^n\}$ be the PoR of flow packets f at node i from the previous source. Previous source of flow f can be initial source or previous cache point of flow f . In fact, φ_i^f means that node i expects to receive transmitted packets from previous source of flow f with the probability of φ_i^f . apparently, the value of φ_0^f is 1 at the source.

As shown in Fig. 1, to calculate the PoR of flow f in the intermediate nodes, the source node of flow f sets φ_0^f to 1 and sends it for node 1. Regarding Eq. 1 each intermediate node $i, i \in \{1, \dots, L_f\}$ calculates the value of φ_i^f as soon as receiving φ_{i-1}^f from node $i-1$, as follow:

$$\varphi_i^f = (1 - PER_{i-1,i}) \cdot \varphi_{i-1}^f, f \in \{1, \dots, F_i^n\} \quad (5)$$

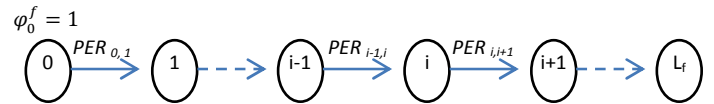


Figure 1. Calculate the PoR of flows in the intermediate nodes.

In DCPS scheme, each intermediate node can be selected as cache point for C_n different flows. The priority of flows in node i , is determined based on their φ_i^f . Thus, if the flow f

passes through the intermediate node i on its way toward the sink and the value of φ_i^f is among C_n flows that have minimum PoR and ($\varphi_i^f < T_\varphi$), then node i is selected as a cache point for flow f . If the node i is selected as a cache point for flow f , before sending φ_i^f to node $i+1$ the value of φ_i^f is set to 1.

As mentioned before, in the DCPS scheme, the intermediate nodes avoid to cache flow packets, the φ_i^f of which is more than T_φ . Because, the packets of flows with high PoR are often received by the sink. So there is no need to retransmit them. Hence, caching them in the intermediate nodes only will lead to the increase of memory consumption at intermediates nodes. Using T_φ in the proposed DCPS scheme will prevent the memory waste at the intermediate nodes.

III. SIMULATION RESULTS

For the evaluation of the DCPS scheme, we implemented DTSN protocol on ns-2 [10]. The considered network topology contains 79 sensor nodes that are located in the form of a tree structure as shown in Fig. 2. Each intermediate node (from node 48 to 77) can be selected as cache point for the flows passing through it. In order to focus on the transport layer performance, it has been assumed that the path of the flows are determined by routing protocol statically.

In considered scenario, every source node (from node 0 to 47) sends a flow containing 500 data packets of 100 bytes each, for the sink. The ACK window size of DTSN protocol was configured as 50 packets. PERs on communication links were applied by random values with 0.1 intervals in the range of 0.0 to 0.5.

In this section, the results of evaluation of DCPS scheme for $C_n = 2$ and $C_n = 3$ are given separately. The value of T_φ is considered 0.8 for both scenarios, too. Furthermore, the results of the evaluation of end-to-end and hop-by-hop retransmission schemes are given for the comparison.

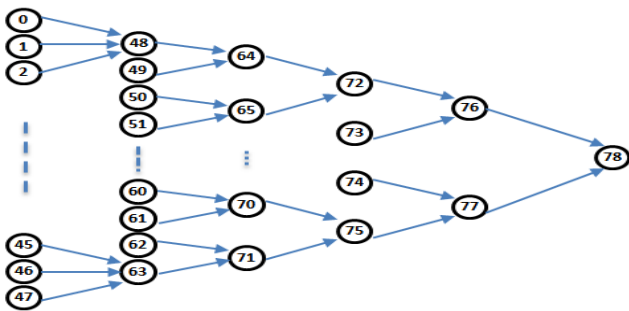


Figure 2. Tree network topology.

A. Transmission Cost

The transmission cost is defined as the average number of required hop-by-hop transmissions for delivering a packet to sink. Fig. 3 shows the transmission cost for different PERs. As expected, the hop-by-hop and end-to-end schemes require minimum and maximum transmission cost, respectively. The

DCPS for both values of C_n has about the same transmission cost with hop-by-hop scheme in the PERs less than 0.3. However, DCPS requires upper transmission cost in the PERs above 0.3.

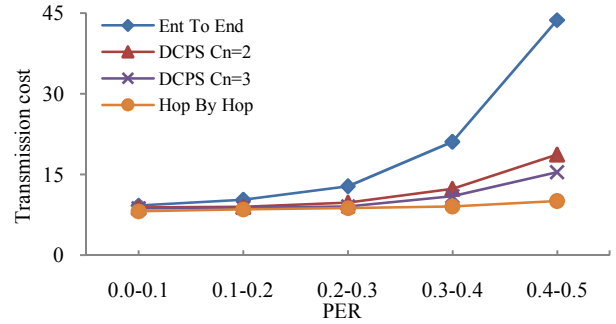


Figure 3. Transmission cost as a function of different PERs.

These results exist in the conditions that as shown in Fig. 2, the intermediate nodes with distance of 1, 2, 3, and 4 levels from sources must cache the packets of 3, 6, 12, and 24 flows in the hop-by-hop retransmission scheme, respectively. But in DCPS with $C_n = 2$ and $C_n = 3$, the intermediate nodes only cache the packets of 2 and 3 flows, respectively. Hence, it can be concluded that increasing the number of cache points for a flow has no effect on decreasing transmission cost in the low PERs.

B. End-to-End Delay

The end-to-end delay for a packet is defined as the time between the initial sending and the time it is received by the sink. Fig. 4 shows this parameter for different PERs.

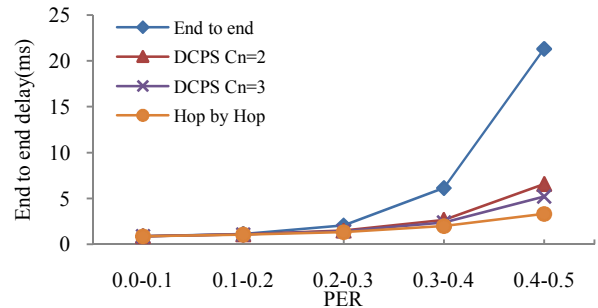


Figure 4. End-to-end delay as a function of different PERs.

Great similarity between Fig. 3 and Fig. 4 demonstrates the existence of the direct relationship between end-to-end delay and transmission cost. Because both parameters depend on the recovery distance. The recovery distance is the number of hops between cache point and loss point. Of course, it should be noted that, end-to-end delay of the packets depends on other parameters such as acknowledgment window.

C. Throughput

Throughput for different PERs are shown in Fig 5. Fig. 5 illustrates, the throughput decreases rapidly with increase in PER. Hence, it seems that the PER has great influence on flow throughput.

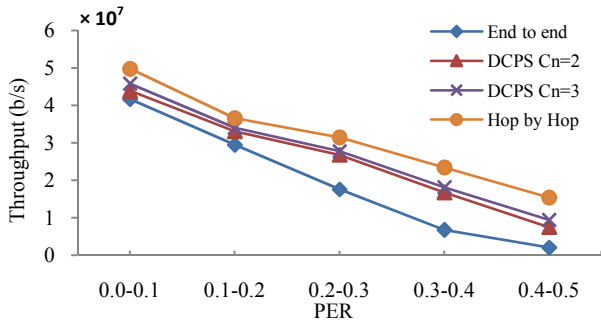


Figure 5. Throughput as a function of different PERs.

The comparison between DCPS and hop-by-hop schemes suggests that on average, the throughput of hop-by-hop scheme is 36% more than DCPS with $C_n = 2$ and 24% more than DCPS with $C_n = 3$. Whereas every flow in hop-by-hop scheme and in the DCPS scheme with $C_n = 2$ and $C_n = 3$, has 5, 1.25, and 1.875 cache points, respectively. The 166% to 300% increase of the number of cache points for each flow is meant by this.

D. Memory Requirements

Fig. 6 shows the maximum and average cache memory required in different PERs for hop-by-hop and DCPS schemes. As shown in Fig. 6, the required cache memory of DCPS is considerably lower than hop-by-hop scheme.

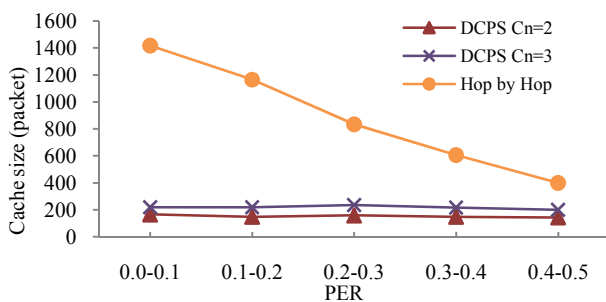


Figure 6. Maximum required cache memory as a function of PERs.

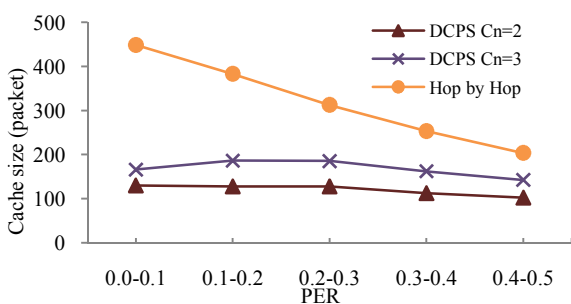


Figure 7. Average required cache memory as a function of PERs.

In the low PERs, the sent packets arrive to intermediate nodes with little loss. However, in high PERs, they are often lost on the communication links. Hence, hop-by-hop scheme consumes maximum cache memory in the low PERs. whereas, in this condition, the PoR of flows is high, so there is no need to recover them. Consequently, cache packets of flows with

high PoR only cause extra cache memory consumption. As mentioned before, in DCPS scheme, an intermediate node caches the packets of the flows whose PoR are less than T_φ . Existence of T_φ in this scheme prevents extra memory consumption.

Load balancing is one of the objectives of management plans whose achievement will be followed by an increase of the lifetime of WSNs. The value of difference between maximum and average required cache memory indicates load balancing on intermediate nodes. The maximum of this value in hop-by-hop scheme is 968 packets whereas, in the DCPS with $C_n = 2$ and $C_n = 3$ are 40 and 57 packets, respectively.

VI. Conclusion and future works

In this paper, using a model based on probability theory, we proposed a distributed cache points selection (DCPS) scheme. Proposed scheme, regarding the memory constraint in sensor nodes, tries to determine the cache points of every flow so that it prevents end-to-end retransmission of loss packets. The proposed scheme is evaluated using ns-2 simulator. Evaluation results show that cache points selection for flows has great impact on efficiency of intermediate caching, memory consumption, and load balancing on intermediate nodes.

With respect to the number and the condition of passing flows, the intermediate nodes experience different conditions. Hence, the algorithms used for them must adapt themselves to different conditions. In the future, we intend to determine the values of T_φ and C_n dynamically based on the conditions of intermediate nodes. Furthermore, proposing a proper policy for cache partitioning where, the cache is partitioned for multiple concurrent flows is considered as a research area.

I. References

- [1] F. Hu and X. Cao, *Wireless Sensor Networks: principles and practice*.
- [2] C. Wan, A. Campbell, and L. Krishnamurthy, "PSFQ: a reliable transport protocol for wireless sensor networks," in *Proc*, 2002, p. 1–11.
- [3] A. Grilo, N. Tigliano, "An analytical model for transport layer caching in wireless sensor networks," *Perform. Eval.*, vol. 69, pp. 227 – 245, Des. 2012.
- [4] B. Marchi, A. Grilo, and M. Nunes, "DTSN: distributed transport for sensor networks," in *Proc*, 2007, p. 165–172.
- [5] A. Dunkels, J. Alonso, T. Voigt, and H. Ritter, "Distributed TCP caching for Wireless Sensor Networks," in *Proc*, 2004.
- [6] C. Wang, M. Daneshmand, B. Li, and K. Sohrawy, "A Survey of Transport Protocols for Wireless Sensor Networks," *IEEE Network*. Vol. 20, pp. 34-40, 2006.
- [7] S. J. Park, R. Sivakumar, I. Akyildiz, and R. Vedantham, "GARUDA: Achieving effective reliability for downstream communication in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 07, pp. 214 –230, 2008.
- [8] F. Stann, and J. Heidemann, "RMST: reliable data transport in sensor networks," in *Proc*, 2003, p. 102–112.
- [9] E. Paek, and R. Govindan, "RCRT: rate-controlled reliable transport for wireless sensor networks," in *Proc*, 2007, p. 305–319.
- [10] The Network Simulator ns-2 website, [Online]. Available: <http://www.isi.edu/nsnam/ns/>.