Many time-sensitive applications impose high requirement on real-time response. There exist many algorithms and routing protocols for efficient data packet delivery. However, previous works set the same retransmission threshold for all the relay nodes along a delivery path. The method decreases the probability of a packet being transmitted through the delivery path within given deadline. In this paper, we focus on finding the optimal retransmission thresholds for the relay nodes, such that the summation of the probability of a packet being transmitted to the next relay node or destination node within the specified deadline is maximized. A distributed greedy algorithm that can be run on sensor node is proposed, which enables the sensor node to adaptively set the optimal retransmission threshold. To avoid dropping the packet forwarded to the destination within given deadline with high probability, we develop a packet dropped protocol based on probabilistic delay bound. Experimental results show that the proposed protocols have better performance.

1. Introduction

Cyber-physical systems (CPS) integrate the computation and physical processes to monitor and control the continuous dynamics of physical world and engineered systems, which revolutionize the approaches for observing the physical world and the process of information access. As extremely critical component in CPS, wireless sensor networks (WSNs) have been envisioned to observe and cognize the complicated physical world at low cost [1, 2]. Deployments of large-scale WSNs have potential to shed light on a variety of monitoring purposes, such as environment monitoring, pollution levels monitoring, and progress of climate change. The discovery of useful knowledge from the massive sensed data is helpful for deeper scientific understanding of human-physical world interaction, which motivates researchers to explore complex and evolving relationships among the data. Plentiful research works have begun focusing on extracting useful knowledge from the Big Data perspective by using the large-scale sensed data.

Large-scale WSNs have been considered as one of the promising, high-value applications for CPS, whose enormous societal impact and economic benefit will be created by the discovery of useful knowledge from the sensed data. Many time-sensitive applications, such as intruder tracking, environment control, and structural health diagnosis impose high requirement on real-time response for decision making and future actions [3–6]. In these applications, deadline misses in data communication may bring about irreparable financial and environmental impacts. Therefore, achieving fast data delivery and response for the massive data generated in WSNs poses new research challenges.

Data delivery delay has been extensively studied in forwarding quality measurement [7–9], sensor network routing, and scheduling [10–13]. Most of the works focus on investigating the metrics to characterize the forwarding quality or minimizing average path delay. Moreover, existing works set the same retransmission threshold for all the sensor nodes along a delivery path. The method is short of taking the link quality and delay requirement into account, which decreases the probability of a packet being successfully transmitted through its delivery path within given deadline. In many cases, the retransmission threshold imposes a significant effect on the probability. To understand the impact of retransmission threshold on the probability of packet delivery
before its deadline, we give an example, which is illustrated in Figure 1.

The number at each link is the probability of a packet being successfully delivered to the next node through the link, denoted by $p_{succ}$, which means that on average it takes $1/p_{succ}$ transmission trials to successfully deliver a packet. We assume that one transmission takes 1 second and the deadline of delivering a packet from node S to node D is 9 s. To deliver the packet to node D before the deadline, the summation of transmission trials along the path must be no more than 9, which is the result of dividing deadline 9 s by one transmission time 1 s. If the retransmission threshold for each link is 2, then it is easily known that the probability of a packet passing the first link is $1 - 0.9 \times 0.9 \times 0.9 = 0.271$. Therefore, the probability of a packet being successfully received by destination node D within its deadline is $0.271 \times 1 \times 1 = 0.271$. However, if the retransmission thresholds for the first link is 6 and the others are 0, then the probability is $(1 - 0.9^6) \times 1 \times 1 \approx 0.522$, which is almost twice as that of the method of setting the same retransmission threshold in advance. The approach of setting the same retransmission threshold did not take the link quality and deadline into account, which decreases the probability of a packet being successfully transmitted through its delivery path within given deadline.

To improve the reliability of packets delivery, the method of packets being transmitted along multiple paths is widely applied. Therefore, in this paper we focus on computing the optimal retransmission thresholds for the relay nodes along a delivery path, such that the summation of the probability of a packet being successfully delivered to the next relay node or destination node within the specified deadline is maximized. To find the optimal retransmission threshold, a distributed greedy algorithm that can be run on sensor node is proposed, which enables the sensor node to adaptively set the optimal retransmission threshold based on the link quality and the remaining time to deadline. Stringent retransmission threshold may make the packet forwarded to the destination within deadline with high probability be dropped. To overcome the shortcoming, we develop an $(\alpha, \beta)$-probabilistic delay bound based packet dropped protocol. The main contributions of the paper are as follows.

(i) The problem of finding optimal retransmission thresholds for the relay nodes along a delivery path is proposed. A formal description of the problem is given, and it can be formalized as an integer optimization problem.

(ii) A distributed greedy algorithm for computing optimal retransmission threshold is proposed and the correctness of this algorithm is proved. The time complexity of the algorithm is $O(\Delta n)$, and the memory complexity is $O(n)$, where $\Delta$ is the given upper bound of the delivery delay and $n$ is the length of the delivery path.

(iii) To avoid dropping the packet forwarded to the destination within the given deadline with high probability, an $(\alpha, \beta)$-probabilistic delay bound based packet dropped protocol is developed.

(iv) Simulation experiments are conducted to evaluate the proposed algorithm and protocol. Experimental results show that our algorithms have better performance in terms of deadline success ratio and real-time ratio.

The rest of this paper is organized as follows. The related works on real-time data delivery are surveyed in Section 2. In Section 3, the optimization problem is described first, and a greedy algorithm for computing optimal retransmission threshold is proposed. Based on Chebyshev inequality, an $(\alpha, \beta)$-probabilistic delay bound based packet dropped protocol is proposed in Section 4. Experimental results are illustrated in Section 5, and Section 6 concludes this paper.

2. Related Works

Data delivery delay is extensively considered in routing metric. The metric of expected transmission count (ETX) was proposed in [14]. ETX measures the expected number of transmissions for successfully delivering a packet over the link, which works well in a homogeneous sigle-radio environment. Effective number of transmissions (ENT) [15] was proposed to enhance ETX by taking the variable link reliability into account. Expected transmission time (ETT) [16] and the path metric of weighted cumulative ETT were proposed for multichannel mesh networks. ETT can be considered as enhanced ETX by counting the hererogeneous channel rate and intraflow interference.

Data delivery delay has also attracted much attention in designing efficient routing protocol. There are three categories of routing protocols that favor the end-to-end delay performance guarantee in WSNs [17]: (I) tree-based routing; (II) optimal routing based on the knowledge of network topology; (III) geographic routing by the information of node position. To shorten the worst-case delay, [18] proposed tree-based routing protocol, which looks up the neighbor table for routing decisions. Tree-based protocol may lead to high energy consumption of the nodes near root of the tree. Generally, geographic routing protocols are not optimal, because they are merely based on one-hop decision [17]. SPEED protocol was proposed in [12], which estimates the forwarding delay. In SPEED protocol, the node of higher relay velocity is selected with higher probability. However, for high efficient energy utilization it is significant to integrate the information effectively.

Motivated by the tradeoff between the energy consumption and real-time routing performance, [19–21] investigated
the optimal construction of virtual backbone and node placement. Aiming at enabling routing with probabilistic delay bounds in wireless sensing and control networks, multitimescale adaptation routing protocol was proposed in [22], which addresses the challenges of dynamic, uncertain link/path delays in real-time routing. To minimize the cost of packet delivery, [23, 24] developed approximation algorithms for the capacitated multicast routing problem.

There are plentiful research works that focus on the analysis of the end-to-end delay in wireless networks. In literature [25], the open queueing network theory was applied to model the WSNs and the analysis for the paths delay based on the model was provided. By mapping the scheduling of real-time periodic data flows in a wireless HART network to real-time multiprocessor scheduling, the analysis of end-to-end delay was given in [26].

However, all the previous works set the same retransmission threshold for the sensor nodes in advance. The method of setting the same retransmission threshold for the sensor nodes along a delivery path is short of taking the link quality and delay requirement into account, which decreases the probability of a packet being successfully transmitted through its delivery path within given deadline. To the best of our knowledge, this paper first investigates the problem of finding optimal retransmission thresholds.

### 3. Greedy Algorithm for Computing Optimal Retransmission Threshold

The method of setting the same retransmission threshold for all sensor nodes has poor performance in real-time data delivery. In this section, we develop a greedy algorithm for computing the optimal retransmission thresholds for the relay nodes along a delivery path, such that the summation of the probability of a packet being successfully transmitted to the next relay node or destination node within the specified deadline is maximized.

#### 3.1. Problem Description

Suppose that the end-to-end path is $P = j_1, j_2, \ldots, j_{n+1}$, where $j_1$ and $j_{n+1}$ are source and destination nodes. Let $p_i$ denote the probability of a transmission failure over the link from $j_i$ to $j_{i+1}$. Then $p_i$ is the probability that a MAC-layer transmission fails due to either collisions or bad channel quality when node $j_i$ forwards a data packet to node $j_{i+1}$. If $K_i$ is the retransmission threshold of node $j_i$, then $1 - p_i^{K_i+1}$ is the probability of a packet being successfully transmitted to node $j_{i+1}$ by node $j_i$. To improve the reliability of packet delivery, packets are transmitted along multiple paths. Therefore, we focus on maximizing the summation of the probability of a packet being successfully transmitted to the next relay node or destination node along the delivery path within the given deadline. Let $\Delta$ denote the end-to-end delay requirement. $L_i$ and $U_i$ are used to denote the given lower and upper bounds of retransmission threshold of node $i$. The problem of finding optimal retransmission thresholds for relay nodes can be formulated as the following integer optimization problem:

\[
\begin{align*}
\text{max} & \quad \sum_{i=1}^{n} (1 - p_i^{K_i+1}) \\
\text{s.t.} & \quad \sum_{i=1}^{n} (K_i + 1) \leq \Delta \\
& \quad L_i \leq K_i \leq U_i, \\
& \quad K_i \in \mathbb{Z}, \quad i \in \{1, 2, \ldots, n\}, \\
& \quad p_i \in (0, 1), \quad L_i, U_i \in (0, +\infty), \quad i \in \{1, 2, \ldots, n\}. \tag{1}
\end{align*}
\]

Obviously, the value of objective function $\sum_{i=1}^{n} (1 - p_i^{K_i+1})$ is maximized if and only if the value of $\sum_{i=1}^{n} p_i^{K_i+1}$ achieves the minimum. For any $i \in \{1, 2, \ldots, n\}$, let $k_i = K_i + 1, l_i = L_i + 1$ and $u_i = U_i + 1$. Then the aforementioned optimization problem is equivalent to the following integer optimization problem:

\[
\begin{align*}
\text{min} & \quad \sum_{i=1}^{n} p_i^{k_i} \\
\text{s.t.} & \quad \sum_{i=1}^{n} k_i \leq \Delta \\
& \quad l_i \leq k_i \leq u_i, \quad k_i \in \mathbb{Z}, \quad i \in \{1, 2, \ldots, n\}, \\
& \quad p_i \in (0, 1), \quad l_i, u_i \in (0, +\infty), \quad i \in \{1, 2, \ldots, n\}. \tag{2}
\end{align*}
\]

Then the problem of finding the optimal retransmission thresholds for each relay node along a delivery path can be defined as follows.

**Input.** (1) The set of the transmission failure probabilities of the nodes along a delivery path, $\{p_1, \ldots, p_n\}$. (2) The end-to-end delivery delay requirement $\Delta$.

**Output.** Optimal retransmission thresholds $k_1', k_2', \ldots, k_n'$, which are optimal solutions for integer optimization problem (2).

Based on the above analysis, the key point is to design an efficient algorithm for finding the optimal solutions to integer optimization problem (2). In the following section, we prove that the optimization problem can be solved by a greedy algorithm. And the correctness of the algorithm is proved.

#### 3.2. Correctness of Greedy Algorithm

**Lemma 1.** For any $i \in \{1, 2, \ldots, n\}$, the following inequalities hold; if $0 < p_i < 1$ and $l_i < u_i$,

\[
p_{i}^{l_{i}+j} - p_{i}^{l_{i}+j-1} < p_{i}^{l_{i}+j+1} - p_{i}^{l_{i}+j} < 0, \tag{3}
\]

where $j \in \{1, 2, \ldots, u_i - l_i\}$.
Proof. Let \( f_i(x) = p_i^x \); it is easily known that \( f''_i(x) > 0 \).

Thus \( f'_i(x) \) is an increasing function with respect to \( x \). According to the differential mean value theorem, the following equalities can be derived:

\[
\begin{align*}
  f_i(l_i + j + 1) - f_i(l_i + j) &= f'_i(\xi) \
  f_i(l_i + j) - f_i(l_i + j - 1) &= f'_i(\eta),
\end{align*}
\]

where \( \xi \in (l_i + j, l_i + j + 1) \) and \( \eta \in (l_i + j - 1, l_i + j) \). Since \( f'_i(x) \) is an increasing function, then \( f'_i(\eta) < f'_i(\xi) \). Then we can derive that \( p_i^{l_i+j} - p_i^{l_i+j-1} < p_i^{l_i+j-1} - p_i^{l_i+j} \). Due to the decreasing of \( f_i(x) \), we have \( p_i^{l_i+j} < p_i^{l_i+j-1} \). In conclusion, \( l_i + \chi_i(l_i), l_2 + \chi_i(l_i), \ldots, l_n + \chi_i(l_i) \) is the optimal solution to problem (7). \( \square \)

Lemma 3. Let \( h = \arg \min_{i \in \{1,2,\ldots,n\}} \{\lambda_i(l_i + 1)\} \); if \( k_i^*, k_1^*, \ldots, k_n^* \) is the optimal solution to the following integer optimization problem, then \( k_i^* \geq l_i + 1 \), where \( \Delta \) is a given positive integer:

\[
\begin{align*}
  \min & \sum_{i=1}^n p_i^{k_i} \\
  \text{s.t.} & \quad \sum_{i=1}^n k_i \leq \Lambda + \sum_{i=1}^n l_i \\
  & \quad l_i \leq k_i \leq u_i, \\
  & \quad k_i \in \mathbb{Z}, \quad i \in \{1,2,\ldots,n\}, \\
  & \quad p_i \in (0,1), \quad l_i, u_i \in (0,\infty), \quad i \in \{1,2,\ldots,n\}. 
\end{align*}
\]

Proof. The proof is by induction on \( \Lambda \). According to Lemma 2, \( k_i^* \geq l_i + 1 \) holds, when \( \Lambda = 1 \). Suppose that the proposition is true, when \( \Lambda < m - 1 \). In the following, we prove that when \( \Lambda = m \), \( k_i^* \geq l_i + 1 \) holds. And the proof is by contradiction.

Suppose that \( k_i^*, k_1^*, \ldots, k_n^* \) is the optimal solution to problem (9), and \( k_i^* < l_i + 1 \). According to the constraints, we know that \( k_i^* = l_i \). For any \( i \in \{1,2,\ldots,n\} \), the objective function is decreasing with respect to \( k_i \). Thus we can derive that \( \sum_{i=1}^n k_i^* = \Lambda + \sum_{i=1}^n l_i \). Then there must exist \( r \in \{1,2,\ldots,n\} \), such that \( k_r^* \geq l_r + 1 \). We define \( k_1^i, k_2^i, \ldots, k_n^i \) as follows:

\[
\begin{align*}
k_i^* & = \left\{ \begin{array}{ll}
k_i^r & i \neq h, \ i \neq r \\
k_i^r + 1 & i = h \\
k_i^r - 1 & i = r.
\end{array} \right.
\end{align*}
\]

It is easily verified that the solution \( k_1^i, k_2^i, \ldots, k_n^i \) satisfies all the constraints; then it is a feasible solution to optimization problem (9). We can derive that \( \sum_{i=1}^n k_i^r - \sum_{i=1}^n l_i = \lambda_i(l_i + 1) + \lambda_i(l_i + 1) = \lambda_i(l_i + 1) + \lambda_i(l_i + 1) = \lambda_i(l_i + 1) \). Since \( \lambda_i(l_i + 1) \) is an increasing function with respect to \( j \), then \( \lambda_i(k_i^r) \geq \lambda_i(l_i + 1) \). Since \( \lambda_i(l_i + 1) = \min_{i \in \{1,2,\ldots,n\}} \{\lambda_i(l_i + 1)\} \), then \( \lambda_i(l_i + 1) > \lambda_i(l_i + 1) \). Therefore, we have that \( \sum_{i=1}^n p_i^{k_i^r} \geq 0 \), which contradicts the fact that \( k_i^*, k_1^*, \ldots, k_n^* \) is the optimal solution. In conclusion, it is proved that \( k_i^* \geq l_i + 1 \). \( \square \)

Theorem 4. If \( k_i^*, k_1^*, \ldots, k_n^* \) is the optimal solution to integer optimization problem (9), then \( k_1^*, k_2^*, \ldots, k_n^* \) is the optimal
solution to the following integer optimization problem, where
\( h = \arg \min_{i \in \{1,2,\ldots,n\}} \{ \lambda_i(l_i + 1) \} \):

\[
\min \sum_{i=1}^{n} p_i k_i
\]

\[
\text{s.t. } \sum_{i=1}^{n} k_i \leq \Lambda - 1 + (l_{i} + 1) + \sum_{i=j \neq h}^{n} l_i
\]

\[
l_i \leq k_i \leq u_i, \quad i \in \{1,2,\ldots,n\} - \{h\}
\]

\[
l_{h} + 1 \leq k_{h} \leq u_{h},
\]

\[
k_i \in \mathbb{Z}, \quad i \in \{1,2,\ldots,n\},
\]

\[
p_i \in (0,1), \quad l_i, u_i \in (0,\infty), \quad i \in \{1,2,\ldots,n\} .
\]

(11)

Proof. Based on Lemma 3, it can be verified that the solution \( k_1^*, \ldots, k_n^* \) satisfies all the constraints of problem (11) and then it is a feasible solution to the problem. In the following, we prove that \( k_1^*, \ldots, k_n^* \) is the optimal solution for Problem (11) by contradiction.

Suppose that \( k_1^*, k_2^*, \ldots, k_n^* \) is the optimal solution to Problem (11), and \( \sum_{i=1}^{n} p_i k_i^* < \sum_{i=1}^{n} p_i k_i^* \). Due to \( l_{h} + 1 \leq k_{h} \leq u_{h} \), we have \( l_{h} \leq k_{h} \leq u_{h} \). Then \( k_1^*, k_2^*, \ldots, k_n^* \) is a feasible solution to integer optimization problem (9). It can be known that \( k_1^*, k_2^*, \ldots, k_n^* \) is a better solution to problem (9), which contradicts the fact that \( k_1^*, \ldots, k_n^* \) is the optimal solution. In conclusion, \( k_1^*, k_2^*, \ldots, k_n^* \) is the optimal solution to Problem (11).

\[\square\]

Theorem 5. Let \( h_r = \arg \min_{i \in \{1,2,\ldots,n\}} \{ \lambda_i(l_i + 1) \} \), for any \( r \in \{2,\ldots,\Lambda\} \), \( h_r \) is defined as follows:

\[
h_r = \arg \min_{i \in \{1,2,\ldots,n\}} \left\{ \lambda_i(j) - \sum_{z=1}^{r-1} \lambda_h \left( l_h + \sum_{m=1}^{\Lambda} \chi_m(h_m) \right) \right\} .
\]

(12)

where \( i \in \{1,2,\ldots,n\} \), \( j \in \{l_1, l_2, \ldots, u_l\} \). Then \( l_1 + \sum_{r=1}^{\Lambda} \chi_r(h_r), l_2 + \sum_{r=1}^{\Lambda} \chi_r(h_r), \ldots, l_n + \sum_{r=1}^{\Lambda} \chi_r(h_r) \) is the optimal solution to integer optimization problem (9).

Proof. The proof is by induction on \( \Lambda \). According to Lemma 2, \( l_1 + \sum_{r=1}^{\Lambda} \chi_r(h_r), l_2 + \sum_{r=1}^{\Lambda} \chi_r(h_r), \ldots, l_n + \sum_{r=1}^{\Lambda} \chi_r(h_r) \) is the optimal solution to problem (9), when \( \Lambda = 1 \). Suppose that when \( \Lambda \leq m - 1 \), the proposition holds. In the following, we prove that when \( \Lambda = m \), the proposition is true.

Suppose that \( k_1^*, k_2^*, \ldots, k_n^* \) is the optimal solution to problem (9), when \( \Lambda = m \). Let \( h = \arg \min_{i \in \{1,2,\ldots,n\}} \{ \lambda_i(l_i + 1) \} \). Based on Theorem 4, we know that \( k_1^*, k_2^*, \ldots, k_n^* \) is the optimal solution to the following integer optimization problem:

\[
\min \sum_{i=1}^{n} p_i k_i
\]

\[
\text{s.t. } \sum_{i=1}^{n} k_i \leq m - 1 + (l_{i} + 1) + \sum_{i=j \neq h}^{n} l_i
\]

\[
l_i \leq k_i \leq u_i, \quad i \in \{1,2,\ldots,n\} - \{h\}
\]

\[
l_{h} + 1 \leq k_{h} \leq u_{h},
\]

\[
k_i \in \mathbb{Z}, \quad i \in \{1,2,\ldots,n\},
\]

\[
p_i \in (0,1), \quad l_i, u_i \in (0,\infty), \quad i \in \{1,2,\ldots,n\} .
\]

(13)

Let \( h_r' = \arg \min_{i \in \{1,2,\ldots,n\}} \{ \lambda_i(j) - \lambda_h(l_{h} + 1) \} \). For any \( r \in \{2,\ldots,m - 1\} \), \( h^*_r \) is defined as follows:

\[
h_r' = \arg \min \left\{ \lambda_i(j) - \lambda_h(l_{h} + 1) \right\}
\]

\[
\left\{ \lambda_h(l_{h} + 1) \right\}
\]

(14)

By the induction hypothesis, \( l_1 + \sum_{r=1}^{\Lambda} \chi_r(h_r), l_2 + \sum_{r=1}^{\Lambda} \chi_r(h_r), \ldots, l_n + \sum_{r=1}^{\Lambda} \chi_r(h_r) \) is the optimal solution to problem (13), and the value of the objective function achieves \( \sum_{i=1}^{n} p_i k_i^* \). Thus we can know that \( l_1 + \sum_{r=1}^{\Lambda} \chi_r(h_r), l_2 + \sum_{r=1}^{\Lambda} \chi_r(h_r), \ldots, l_n + \sum_{r=1}^{\Lambda} \chi_r(h_r) \) is the optimal solution to problem (9), when \( \Lambda = m \).

Let \( h_1 = \arg \min_{i \in \{1,2,\ldots,n\}} \{ \lambda_i(l_i + 1) \} \), and, for any \( r \in \{2,\ldots,m\} \), \( h_r \) is defined based on formula (12). It is easily known that for any \( i \in \{1,2,\ldots,n\} \), \( l_1 + \sum_{r=1}^{\Lambda} \chi_r(h_r), \chi(h) = l_r + \sum_{r=1}^{\Lambda} \chi_r(h_r) \) is the optimal solution to problem (9).

\[\square\]

3.3. Greedy Algorithm for Computing Optimal Retransmission Threshold. From Theorems 4 and 5, we know that integer optimization problem (9) exhibits the optimal substructure and greedy-choice properties, which are two key ingredients of the greedy approach. In this section, a greedy algorithm for computing the optimal retransmission threshold is proposed.

Suppose that the end-to-end path is \( P = j_1, j_2, \ldots, j_m \), where \( j_1 \) and \( j_{m+1} \) are source and destination nodes, respectively. Let \( p_i \) denote the transmission failure probability over the link from \( j_i \) to \( j_{i+1} \). Suppose that relay node \( j_i \) receives a data packet with delay requirement \( \Delta' \). The greedy algorithm for computing optimal retransmission threshold runs at sensor node \( j_i \) as follows. The retransmission thresholds for nodes \( j_0, j_1, \ldots, j_m \) are initialized as \( L_0, L_{j_1}, \ldots, L_{j_m} \), respectively, where \( L_h \) is the given lower bound of retransmission threshold of node \( j_h \). First, for each \( h \in \{i, i+1, \ldots, n\} \), compute \( p_i \). Then, the \( L_h \) with the minimum value \( L_{j_{i+1}} - p_i \) is incremented by one. The above procedures are repeated \( \Delta' - \sum_{i=1}^{n} (L_i + 1) \) times. The detail of the algorithm is described in Algorithm 1.

For a given delay requirement \( \Delta \) and a delivery path consisting of \( n \) sensor nodes, the cost for computing the optimal retransmission threshold is \( O(n\Delta) \), and the memory complexity is \( O(n) \).
Input: $\Delta', \{p_i, p_{i+1}, \ldots, p_n\}, \{L_i, L_{i+1}, \ldots, L_n\}, \{U_i, U_{i+1}, \ldots, U_n\}$  
Output: optimal retransmission threshold of node $i$  

1. $\Delta' = \Delta - \sum_{h=i}^{n} (L_h + 1)$,  
2. for $k = 1$ to $\Delta'$ do  
3. $\text{term} = 0$  
4. for $h = i$ to $n$ do  
5. if $L_h + 1 \leq U_h$ then  
6. $\lambda_h = p_h^{L_h+1} - p_h^{L_h+2}$;  
7. else  
8. $\lambda_h = \infty$;  
9. if $\lambda_h < \text{term}$ then  
10. $\text{term} = \lambda_h$;  
11. $\kappa = \text{arg}\{\text{term}\}$;  
12. $L_\kappa = L_\kappa + 1$;  
13. return $L_i$

4. $(\alpha, \beta)$-Probabilistic Delay Bound Based Packet Dropped Protocol

In previous section, we investigate the problem of finding the optimal retransmission thresholds for the relay nodes along a given end-to-end delivery path. Due to the uncertainty of wireless link, stringent retransmission threshold may make the packet forwarded to the destination within deadline with high probability be dropped. Since computing the probabilistic distribution of path delays is NP-hard [27], then we develop an $(\alpha, \beta)$-probabilistic delay bound based packet dropped protocol in this section.

The main idea of the proposed protocol is as follows. When a packet fails to be transmitted after a predefined number of retransmissions, if the upper bound of the probability of the delivery delay within the deadline is smaller than $\alpha$, then the packet is dropped; if the upper bound of the probability of the delivery delay exceeding the deadline is smaller than $\beta$, then the packet is transmitted once again. In the following section, the probabilistic path delay bounds are derived.

4.1. Probabilistic Path Delay Bound. Given a path $P = j_1, j_2, \ldots, j_{n+1}$ from source node $j_1$ to destination node $j_{n+1}$, let $d_{j_i}$ denote the number of transmission trials when $j_i$ forwards the packet to node $j_{i+1}$. Then the end-to-end delivery delay over the path is $d_P = \sum_{i=1}^{n} d_{j_i}$, where $d_{j_i}$ follows geometric distribution $G(1 - p_i)$. Since per-packet transmission time is stable and uncorrelated [22], we have that $E[d_P] = \sum_{i=1}^{n} (1 - p_i)^{-1}$, $\text{Var}[d_P] = \sum_{i=1}^{n} p_i(1 - p_i)^{-2}$.

Property 1. Suppose that $\Delta$ is the delay requirement, let $\Delta = \Delta - E[d_P]$. If $\Delta < E[d_P]$, then the following inequality holds:

$$
\text{Pr}\left(d_P \leq \Delta\right) \leq \frac{\text{Var}[d_P]}{\text{Var}[d_P] + \delta^2}.
$$

If $\Delta \geq E[d_P]$, then the following inequality can be known:

$$
\text{Pr}\left(d_P \geq \Delta\right) \leq \frac{\text{Var}[d_P]}{\text{Var}[d_P] + \delta^2}.
$$

Proof. According to one-sided Chebyshev inequality, the following inequality holds, if $t < 0$ and $\text{Var}[x]$ is bounded:

$$
\text{Pr}\left(x \leq E(x) + t\right) \leq \frac{\text{Var}[x]}{\text{Var}[x] + t^2}.
$$

Similarly, the following inequality holds, if $t \geq 0$ and $\text{Var}[x]$ is bounded:

$$
\text{Pr}\left(x \geq E(x) + t\right) \leq \frac{\text{Var}[x]}{\text{Var}[x] + t^2}.
$$

Thus formula (15) and (16) are easily obtained.

4.2. $(\alpha, \beta)$-Probabilistic Delay Bound Based Packet Dropped Protocol. Suppose that relay node $j_i$ fails to transmit a data packet to node $j_{i+1}$ after a predefined number of retransmission trials, and the remaining time to the deadline is $\Delta'$. For given probability thresholds $\alpha$ and $\beta$, $(\alpha, \beta)$-probabilistic delay bound based packet dropped protocol implements at sensor node $j_i$ as follows.

Firstly, node $j_i$ computes the expectation and variance of the delivery delay over the path $j_{i+1}, j_{i+2}, \ldots, j_n$, denoted by $\mu$ and $\sigma$, respectively. Let $\delta = \Delta' - 1 - \mu$ be the difference between $\Delta' - 1$ and the expectation of the delivery delay.

Secondly, the node decides whether the packet will be dropped as follows.

Case 1 ($\Delta' - 1 < \mu$). If $\sigma/(\sigma + \delta^2) \leq \alpha$, then the packet is dropped; otherwise the packet is transmitted once again.

Case 2 ($\Delta' - 1 \geq \mu$). If $\sigma/(\sigma + \delta^2) \leq \beta$, then the packet is transmitted once again; otherwise the packet is dropped.

The detail of the protocol is described in Algorithm 2.
Algorithm 2: \((\alpha, \beta)\)-Probabilistic delay bound based packet dropped protocol.

5. Performance Evaluation

In this section, a series of experiments have been conducted to evaluate the performance of the retransmission threshold setting protocol based on greedy algorithm for computing optimal retransmission threshold (GAORT) and probabilistic delay bound based packet dropped (PDBPD) protocol. The simulations are carried out by MATLAB.

100 sensor nodes and the sink are randomly deployed into a region of size 200 m \(\times\) 200 m (m for meters) and all the sensors have the same transmission radius 25 m. In each simulation, source node and destination node are randomly selected. Each simulation is repeated 100 times and the simulation result corresponds to the average value over 100 times. To understand the benefits of the GAORT based protocol, the comparison with General Method is conducted. The main idea of General Method is that we set the same retransmission threshold for all the sensor nodes along a delivery path. In the experimental simulation, \(\Delta\) is the result of the deadline divided by one transmission time.

The first group of experiments is to investigate the deadline success ratio (DSR) of the GAORT based protocol. In this paper, deadline success ratio is the percentage of the packets delivered to the destination before given deadlines among all the packets. And DSR can be calculated as follows:

\[
\text{DSR} = \frac{\# \text{of packets delivered to the destination before given deadlines}}{\# \text{of all of the packets}}. \tag{19}
\]

Figure 2(a) shows the comparison of DSR between the GAORT based protocol and General Method, when the length of delivery path is 10 and the probability of transmission failure, \(p_i\), follows normal distribution \(N(0.4, 0.2)\). The GAORT based protocol has better performance in terms of deadline success ratio. For example, when the deadline is 0.11 s, the DSRs of the GAORT based protocol and General Method are 82% and 69%, respectively. Figure 2(b) depicts the relationship between the DSRs and the deadlines, when the probability of transmission failure, \(p_i\), follows uniform distribution \(U[0.4, 0.9]\). Figure 2(c) demonstrates the comparison of DSRs under different lengths of delivery path. As expected, the GAORT based protocol can achieve a higher deadline success ratio, since delay requirement and link quality have been taken into account by our protocol. Experimental results show that the GAORT based protocol can optimize the probability of packet delivery within given deadline and the proposed protocol increases the number of retransmissions over the link of inferior quality.

The second group of experiments is to investigate the real-time ratio (RTR) of the GAORT based protocol. Real-time ratio is the percentage of the packets delivered to the destination before given deadlines among the packets successfully delivered to the destination. And RTR can be calculated as follows:

\[
\text{RTR} = \left(\frac{\# \text{of packets delivered to the destination before given deadlines}}{\# \text{of packets delivered to the destination}}\right) \times \left(\frac{\# \text{of all of the packets}}{\# \text{of packets delivered to the destination before given deadlines}}\right)^{-1}. \tag{20}
\]

As shown in Figure 3(a), the GAORT based protocol can reach higher real-time ratio than that of General Method,
when the length of delivery path is 10 and the probability of transmission failure, $p_i$, follows normal distribution $N(0.4, 0.2)$. Figure 3(b) demonstrates the comparison of RTR, when the probability of transmission failure, $p_i$, follows uniform distribution $U[0.4, 0.9]$. Figure 3(c) shows that the GAORT based protocol has better performance in terms of RTR, when the length of delivery path is increased from 5 to 10. For example, the RTR of the GAORT based protocol is more than 75%. Experimental results indicate that the protocol based on GAORT can reduce the number of retransmission for the packets, which are forwarded to the destination nodes before given deadlines with low probabilities. Therefore, the GAORT based protocol improves the real-time ratio and the energy efficiency.

The third group of experiments is to investigate the performance of PDBPD protocol. The comparisons between Normal Method and PDBPD protocol are carried out. The main idea of Normal Method is that a data packet is dropped after a predefined number of retransmissions, where the retransmission threshold is the result by implementing GAORT. As shown in Figures 4(a) and 4(b), PDBPD protocol achieves higher deadline success ratio than that of Normal Method, when the probability of transmission failure, $p_i$,
follows normal distribution $N[0.6, 0.2]$. For example, the DSRs of PDBPD protocol and Normal Method are 80% and 73%, respectively, when the deadline is 0.12 s and the length of delivery path is 10. Figure 4(c) demonstrates the comparison of DSRs, when the probability of transmission failure follows uniform distribution $U[0.3, 0.6]$. Experimental results show that PDBPD protocol can avoid dropping the packets, which are forwarded to the destination before given deadlines with high probabilities.

6. Conclusion

As one of the promising, high-value applications for CPS, deployments of large-scale WSNs have potential to shed light on a variety of monitoring purposes. Many time-sensitive applications impose high requirement on real-time response. There exist many algorithms and routing protocols for efficient data packet delivery. However, previous works set the same retransmission threshold for all the relay nodes along a delivery path. The method decreases the probability of a packet being transmitted through the delivery path within given deadline. In this paper, we focus on finding the optimal retransmission thresholds for the relay nodes, such that the summation of the probability of a packet being transmitted to the next relay node or destination node within the specified deadline is maximized. A distributed greedy algorithm that can be run on sensor node is proposed, which enables the sensor node to adaptively set the optimal retransmission

![Figure 3: Real-time ratio of GAORT based protocol.](image-url)
threshold. To avoid dropping the packet forwarded to the destination within given deadline with high probability, we develop an \((\alpha, \beta)\)-probabilistic delay bound based packet dropped protocol. Experimental results show that the proposed protocols have better performance.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

This work is supported in part by the National Grand Fundamental Research 973 Program of China under Grant 2012CB316200, the Key Program of the National Natural Science Foundation of China under Grants 61033015 and 60953001, and the Major Program of National Natural Science Foundation of China under Grant 61190115.

**References**

[1] J. Li, S. Cheng, H. Gao, and Z. Cai, “Approximate physical world reconstruction algorithms in sensor networks,” *IEEE Transactions on Parallel and Distributed Systems*, 2014.

[2] X. Cheng, A. Thaeler, G. Xue, and D. Chen, “TPS: A time-based positioning scheme for outdoor wireless sensor networks,” in *Proceedings of the IEEE 23rd Conference on Computer Communications (INFOCOM ’04)*, vol. 4, pp. 2685–2696, March 2004.
[3] Y. Li, C. S. Chen, Y.-Q. Song, Z. Wang, and Y. Sun, “Enhancing real-time delivery in wireless sensor networks with two-hop information,” IEEE Transactions on Industrial Informatics, vol. 5, no. 2, pp. 113–122, 2009.

[4] A. Kadri, E. Yaacoub, M. Mushtaha, and A. Abu-Dayya, “Wireless sensor network for real-time air pollution monitoring,” in Proceedings of the 1st International Conference on Communications, Signal Processing and Their Applications (ICCSPA ’13), pp. 1–5, IEEE, February 2013.

[5] G. Liu, R. Tan, R. Zhou, G. Xing, W.-Z. Song, and J. M. Lees, “Volcanic earthquake timing using wireless sensor networks,” in Proceedings of the 12th International Conference on Information Processing in Sensor Networks (IPSN ’13), pp. 91–102, ACM, April 2013.

[6] M. Ding, D. Chen, K. Xing, and X. Cheng, “Localized fault-tolerant event boundary detection in sensor networks,” Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM ’05), 2005.

[7] J. Wang, Y. Liu, M. Li, W. Dong, and Y. He, “QoF: towards comprehensive path quality measurement in wireless sensor networks,” in Proceedings of the IEEE INFOCOM, pp. 775–783, Shanghai, China, April 2011.

[8] S. Lin, G. Zhou, and K. Whitehouse, “Towards stable network performance in wireless sensor networks,” in Proceedings of the 30th IEEE Real-Time Systems Symposium (RTSS ’09), pp. 227–237, Washington, DC, USA, 2009.

[9] H. Zhang, L. Sang, and A. Arora, “Comparison of data-driven link estimation methods in low-power wireless networks,” IEEE Transactions on Mobile Computing, vol. 9, no. 11, pp. 1634–1648, 2010.

[10] H. Li, Y. Cheng, and C. Zhou, “Minimizing end-to-end delay: a novel routing metric for multi-radio wireless mesh networks,” in Proceedings of the 28th IEEE International Conference on Computer Communications (INFOCOM ’09), pp. 46–53, Rio de Janeiro, Brazil, April 2009.

[11] S. Yin, Y. Xiong, Q. Zhang, and X. Lin, “Traffic-aware routing for real-time communications in wireless multi-hop networks,” Wireless Communications and Mobile Computing, vol. 6, no. 6, pp. 825–843, 2006.

[12] T. He, J. A. Stankovic, C. Lu, and T. Abdelzaher, “SPEED: a stateless protocol for real-time communication in sensor networks,” in Proceedings of the 23rd IEEE International Conference on Distributed Computing Systems, pp. 46–55, Providence, RI, USA, May 2003.

[13] L. Guo, Y. Li, and Z. Cai, “Minimum-latency aggregation scheduling in wireless sensor network,” Journal of Combinatorial Optimization, 2014.

[14] D. S. J. de Couto, D. Aguayo, J. Bicket, and R. Morris, “A high-throughput path metric for multi-hop wireless routing,” in Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom ’03), pp. 134–142, San Diego, Calif, USA, September 2003.

[15] C. E. Koksal and H. Balakrishnan, “Quality-aware routing metrics for time-varying wireless mesh networks,” IEEE Journal on Selected Areas in Communications, vol. 24, no. 11, pp. 1984–1994, 2006.

[16] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh networks,” in Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom ’04), pp. 114–128, October 2004.

[17] Y. Li, C. S. Chen, Y.-Q. Song, Z. Wang, and Y. Sun, “Enhancing real-time delivery in wireless sensor networks with two-hop information,” IEEE Transactions on Industrial Informatics, vol. 5, no. 2, pp. 113–122, 2009.

[18] B. Nefzi and Y.-Q. Song, “Performance analysis and improvement of ZigBee routing protocol,” in Proceedings of the 7th IFAC International Conference on Fieldbuses and Networks in Industrial and Embedded Systems (FeT ’07), pp. 199–206, November 2007.

[19] A. Boukerche, X. Cheng, and J. Linus, “Energy-aware data-centric routing in microsensor networks,” in Proceedings of the 6th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM ’03), pp. 42–49, September 2003.

[20] X. Cheng, X. Huang, D. Li, W. Wu, and D.-Z. Du, “A polynomial-time approximation scheme for the minimum-connected dominating set in ad hoc wireless networks,” Networks, vol. 42, no. 4, pp. 202–208, 2003.

[21] X. Cheng, D.-Z. Du, L. Wang, and B. Xu, “Relay sensor placement in wireless sensor networks,” Wireless Networks, vol. 14, no. 3, pp. 347–355, 2008.

[22] X. Liu, H. Zhang, Q. Xiang, X. Che, and X. Ju, “Taming uncertainties in real-time routing for wireless networked sensing and control,” in Proceedings of the 13th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc ’12), pp. 75–84, June 2012.

[23] Z. Cai, G. Lin, and G. Xue, “Improved approximation algorithms for the capacitated multicast routing problem,” in Proceedings of the 11th Annual International Conference (COCOON ’05), pp. 136–145, Kunming, China, August 2005.

[24] Z. Cai, R. Goebel, and G. Lin, “Size-constrained tree partitioning: approximating the multicast k-tree routing problem,” Theoretical Computer Science, vol. 412, no. 3, pp. 240–245, 2011.

[25] T. Qiu, F. Xia, L. Feng, G. Wu, and B. Jin, “Queueing theory-based path delay analysis of wireless sensor networks,” Advances in Electrical and Computer Engineering, vol. 11, no. 2, pp. 3–8, 2011.

[26] A. Saifullah, Y. Xu, C. Lu, and Y. Chen, “End-to-end delay analysis for fixed priority scheduling in WirelessHART networks,” in Proceedings of the 17th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS ’11), pp. 13–22, April 2011.

[27] R. A. Guérin and A. Ordóñez, “QoS routing in networks with inaccurate information: theory and algorithms,” IEEE/ACM Transactions on Networking, vol. 7, no. 3, pp. 350–364, 1999.
