A Reconfigurable Routing Protocol for Free Space Optical Sensor Network

Rong Xie, Fei Tong, Dongki Kang, Young-Chon Kim*
Department of Computer Engineering, Chonbuk National University, Jeonju, Korea
Email: {xierong,feitong,dkkang}@jbnu.ac.kr
*Corresponding author: yckim@chonbuk.ac.kr

Abstract—Recently, there is a growing interest in free space optical sensor network (FSOSN) for wireless applications due to the advantages of lower energy consumption, lower cost, and larger capacity. However, the unidirectional communication of FSOSN suffers from a dramatic decrease of connectivity if certain nodes die of the depletion of energy, which further results in huge loss of packets in data transmission. Thus, this paper proposes a reconfigurable routing protocol for FSOSN to overcome this problem by reconfiguring the network virtual topology. The protocol is operated in four steps: 1) virtual topology construction, 2) routing establishment, 3) routing table update, 4) reconfigurable routing. During data transmission, the data is firstly routed through the shortest hop path. Then the reconfiguration is initiated by the node whose residual energy is below a threshold. The nodes affected by this dying node are classified into two types: maintenance nodes and adjustment nodes. They are reconfigured according to their types. An energy model is set up and the performance of the protocol is evaluated and compared with simple-link, a modified secure integrated routing and localization scheme, and a direct reconfiguration scheme in terms of connectivity, the number of living node and packet delivery ratio by OPNET simulation.

I. INTRODUCTION

In the past decade, wireless sensor network (WSN), which consists of many small sensor nodes that communicate via radio frequency (RF), has been introduced as a new technology of wireless network. It has extensive applications, especially in area monitoring where sensor nodes collect the interested data or just sense the presence or the absence of an interested event in the field. But this RF-based sensor network suffers from potential issues such as signal interference, attenuation and collision. As one of the promising candidates, free space optical sensor network (FSOSN) based on free space optical transmission has shown its merits of lower energy consumption, lower cost and larger capacity over the traditional RF-based sensor network [1],[2],[3],[4].

However, the performance of FSOSN is restricted to its unidirectional communication in which a node may directly talk to its neighbors but be unable to hear from these neighbors. This characteristic requires efficient routing protocols for data transmission. There has been several recent works on routing in FSOSN. In [5], J. Diaz et al. proposed simple-bro and simple-link protocols to ensure the broadcast from the base station (BS) to each mote and the communication initiated from each mote to the BS respectively. To support the communication among sensor nodes and that between the BS and each sensor node, two efficient routing algorithms, namely neighborhood discovery algorithm (NDA) and base station discovery algorithm (BDA), were proposed in [6]. Exploiting the security benefits of link directionality, U. Ndili Okorafor et al. in [7] proposed a secure integrated routing and localization scheme (SIRLoS) to address security breaches and attacks. In [8], the performance of the SIRLoS was well evaluated.

There is still a great deal of research focused on reconfiguring sensors during routing in FSOSN. The rest of the paper continues with the following structure. Section II presents our preliminaries. Section III introduces the detail of the RRP. Sections IV and V give the simulations and performance analysis respectively. Finally, section VI concludes the paper.

II. PRELIMINARIES

In this section, the node architecture of our work and the network setup are introduced.

A. Node Architecture

Currently, there is a great deal of research focused on free-space optics based sensor network system [10],[11],[12],[13]. Different assumptions of the node architecture and beam characteristics will affect the advantages of different protocols. In our work, we assume a simple node architecture similar to that proposed in [10], which consists of a sensor, a control
unit, an active transmitter, a passive transmitter with corner-cube retro reflector (CCR), a receiver with photo detector, and a battery. Fig. 1 shows the architecture.

B. Network Setup

The network is composed of a set of nodes that are randomly distributed in a given area. Each sensor node with random position \((x_i, y_i)\) has a given communication radius \(r\) and a random direction \(\theta_i\). It can orient the active transmitting laser within a scanning area that covers a contiguous sector of \(\alpha\) degrees. This angle of the scanning area is called scanning angle and its range is \([-\frac{\alpha}{2} + \theta_i, \frac{\alpha}{2} + \theta_i]\). The receiver is omni-directional and can receive data from any direction. In Fig. 2(a), for node \(S_i\) to hear \(S_j\), we must have that: \(d(S_i, S_j) \leq r\) and \((x_j, y_j) \in \Phi_i\), where \(d(S_i, S_j)\) is the Euclidean distance between node \(i\) and \(j\), and \(\Phi_i\) is the scanning area of node \(i\). In this case, \(S_i\) may directly talk to \(S_j\); However, \(S_j\) can only talk to \(S_i\) via a multi-hop with other nodes acting as routers. Following the definition in [6], \(S_i\)’s forward neighborhood, denoted by \(F Neb(S_i)\), is the set of all nodes that \(S_i\) can talk to directly. Similarly, \(S_i\)’s backward neighborhood, denoted by \(B Neb(S_i)\), is a set of all nodes that can talk to \(S_i\) directly. The nodes in \(F Neb(S_i)\) are called \(S_i\)’s predecessors and the nodes in \(B Neb(S_i)\) are called \(S_i\)’s successors. In Fig. 2(b), \(S_j\) is a successor of \(S_i\) and \(S_i\) is a predecessor of \(S_j\).

The hierarchical network is set up through an interrogation process in which the BS sends interrogating beam of light to the entire network. Any node that shares a direct line-of-sight (LOS) path with the BS can establish bi-directional communication with the BS using its CCR. Such nodes are called cluster heads (CHs). Other sensors, which have to route multi-hops to get to the CHs for contacting with the BS, are general sensors (GSs). They transmit data by the active transmitter, which consumes larger amount of energy than that by the CCR. Fig. 2(c) shows a simple topology of a FSOSN and Fig. 2(d) is its hierarchical topology. Note that the CHs \((S_1, S_3, S_{10})\) communicate with the BS bi-directionally.

III. RECONFIGURABLE ROUTING PROTOCOL

In this section, we introduce the details of the RRP. It is operated in four steps, namely virtual topology construction, routing establishment, routing table update, and reconfigurable routing.

A. Virtual Topology Construction

This step is to construct the virtual topology of the network for the BS by flooding a circuit discovery packet (CDP) initiated from the BS. The process is similar to BDA in [6], but we use CDP in our protocol. The CDP contains a hops-traversed (HT) field and a BS-information (BSI) field in its header. The HT counts the number of hops it has traversed and the BSI records the BS location information. When a CH receives a CDP, it has to verify this CDP by checking the HT. If HT = 1, it means the CDP does not travel through the GSs and contains no useful information. Otherwise the CH increases the HT by one, and appends information including its ID and location to the CDP. Then it forwards the CDP to its successors in the down-link case (HT = 1), or to the BS in the up-link case (HT > 1). When a GS receives a CDP, it records the BSI contained in the packet, increases the HT by one and then forwards the CDP to its successors. Note that the CDP is expired if the HT is larger than a given constant \(\delta\). When a CDP returns to the BS, the BS gets the nodes’ location and extracts the path which is formed by the ID sequence in the CDP to make entries into a global graph \((n \times n)\). The pseudo-code of this step is shown in Algorithm 1.

Algorithm 1 Global Graph Construction

| Step | Description |
|------|-------------|
| 1. BS → CHs : CDP \([HT = 0, \text{BSI}]\) |
| 2. CH: |
| while \(HT < \delta\) do |
| if \(HT = 1\) then |
| discards the CDP |
| else |
| \(HT ← HT + 1\) |
| appends \((ID_{CH_i}, (x_{CH_i}, y_{CH_i}))\) to the CDP |
| if \(HT = 1\) then |
| CH → Successors (CH) : CDP |
| else |
| CH → BS : CDP |
| end if |
| end if |
| end while |
| 3. GS: |
| while \(HT < \delta\) do |
| records the BSI |
| \(HT ← HT + 1\) |
| appends \((ID_{GS_i}, (x_{GS_i}, y_{GS_i}))\) to the CDP |
| GS → Successors (GS) : CDP |
| end while |

B. Routing Establishment

In this step, the BS establishes routes for each node by firstly constructing an up-link graph \((n \times n)\) that includes all of the shortest hop paths from the nodes to it. Then it extracts the routing table from this up-link graph for each node. Fig. 4 is the format of a GS’s routing table. Two fields are included:
the next hop set and the cost. The next hop set contains the IDs of the next hop candidates for current node to get to the BS. The cost is the number of hops from current node to the BS.

Algorithm 2 shows the pseudo-code of this step. The construction of the up-link graph is similar to the level-first search (LFS) algorithm. Two queues, $P\_\text{Queue}$ and $H\_\text{Queue}$, are used in pairs. $P\_\text{Queue}$ is for queuing the nodes from low level to high level, and $H\_\text{Queue}$ counts the number of hops from corresponding node in $P\_\text{Queue}$ to the BS. Note that the CH is in the lowest level, which only takes one hop to the BS. The BS maintains a $Hop$ value for each node that records the smallest number of hops from the sensor node to the BS. Fig. 5 shows the up-link graph for the simple FSOSN in Fig. 2(c) and Tab. I gives the results of the routing table of each GS.

### C. Routing Table Update

After establishing the routes, the BS uni-casts the computed routing table to each node. Upon receipt, the node stores the routing table for data transmission.

### D. Reconfigurable Routing

When each node gets its routing table, the data transmission begins. Initially, the node forwards data to the BS by randomly selecting one next hop in its routing table. As time goes by, if the remaining energy of a node reaches the threshold, it immediately sends an SOS message to the BS for reconfiguration. This threshold $E_{\text{threshold}}$ is set by the following equation:

$$E_{\text{threshold}} = p \times E_{\text{init}}$$  \hspace{1cm} (1)

where $E_{\text{init}}$ is the initial energy of a node and $p$ is a percentage that can be appropriately set by the network designer according to the specific traffic load and application.

When the BS receives this SOS message, it figures out those nodes that may be affected by this dying node (DN). These affected nodes can be classified into two types: maintenance nodes (MNs) and adjustment nodes (ANs). MNs are nodes whose routing tables need to be modified, and ANs are nodes that need orientation and transmission range adjustment for communication. Both MNs and ANs are the predecessors of the DN, yet the difference is that MNs can get to the BS through other nodes while the ANs can only communicate with the BS through the DN. In Fig. 5, $S_4$ is a MN and $S_6$ is an AN when $S_7$ is dying. After the classification, the BS reconfigures the DN, MNs and ANs as follows:

For the DN, the BS sends it a dying notification message. Then the DN sets up a bi-directional path between itself and the BS. If the DN is a CH, the bi-directional path has already been set up by the CCR; otherwise, if the DN is a GS, this can be done by re-orienting the DN’s active transmitter and adjusting its transmission range to reach the BS. After the bi-directional path establishment, the DN is isolated from others, and only transmits its own packet to the BS directly.

For all the MNs, as the path to the DN becomes invalid, the BS sends each of them an routing maintenance message. Upon receipt, each MN eliminates the ID of the DN in the next hop set of its routing table to exclude this invalid path.

For all the ANs, as the only route to the BS is through the DN, orientation and transmission range adjustment are needed. Firstly, the BS computes a new orientation for each AN, then it sends the ANs messages for reconfiguration. To depict this
scheme, some useful denotation is introduced in Tab. II. Note that we define a circle area for each AN as $A_{AN}$, which is the circle centrally located at the AN with a radius of transmission range. Fig. 6 shows the flow chart of reconfiguring the AN. At first, the AN tries to find a living CH in its circle area. If such a CH exists, then it adjusts its orientation to this CH. If there are more than one qualified CHs, one of them is selected randomly. However, in the case that no CH is available, the AN tries to find a GS that satisfies connectivity requirement for re-orientation. If there are two or more candidates, then the one with the smallest number of hops to the BS is selected. If none of the qualified CH and GS is found in $A_{AN}$, then the AN becomes an up cluster that should adjust its orientation and transmission range to reach the BS for up-link communication. Note that a GS satisfies the connectivity requirement means that there is at least one path available for it to communicate with the BS.

IV. SIMULATIONS

The simulation is conducted through OPNET Modeler 10.0. In this section, an energy model is firstly introduced, followed by a simple description of protocols for comparison. Then we give the simulation scenario and evaluation metrics.

A. Energy Model

According to the node architecture in Fig. 1, the total energy required for transmitting ($E_{TX}$) includes the energy consumed by the active transmitter ($E_{AX}$) and that consumed by the passive transmitter ($E_{PX}$):

\[ E_{TX} = E_{AX}(m, d, \varphi) + E_{PX}(n) \]  (2)

For simplicity, we assume the energy for the active transmitter to complete a full scan is the same with that to complete a transmission fixating on a specific direction. This energy loss is proportional to the laser scanning area and a coefficient $\varepsilon = 4 \text{ pJ/bit/m}^2$ is needed for the active transmitter to achieve an acceptable $\frac{E_{TX}}{E_{AX}}$. During the data transmission, if the number of bits transmitted by the active transmitter is $m$, the scanning angle is $\varphi$, the transmission range is $d$, then the energy for the active transmitter is:

\[ E_{AX}(m, d, \varphi) = \varepsilon \times \varphi \times d^2 \times m \]  (3)

If the total number of bits transmitted by the CCR is $n$, and the transmitted energy per bit for the CCR is $E_{CCR}$, then the energy for passive transmitter is:

\[ E_{PX}(n) = E_{CCR} \times n \]  (4)

For the receiver, a received energy per bit is $E_R$, and the total energy for receiving $k$ bits is:

\[ E_{RX} = E_R \times k \]  (5)

The parameters of the Energy Model is shown in Tab. III and it’s slightly better than that in [11].

B. Protocols for Comparison

The performance of the proposed protocol is compared with the simple-link, a modified SIRLoS (M-SIRLoS) and a direct reconfiguration scheme (DRS). The simple-link is specified in [5]. The M-SIRLoS uses the shortest hop path routing strategy in which each sensor node always chooses the path of the lowest cost in its successor routing table (SRT) [8] during up-link communication. The DRS is a simple strategy that each GS immediately adjusts its orientation and transmission range to the BS once it gets the BSI.

C. Simulation Scenario

The network is consisted of 100 sensor nodes that are scattered in a 100$\times$100 m$^2$ area with network connectivity. This means firstly each node has at least one path to get to the BS. 10% of the nodes are CHs. We consider a simple scenario in monitoring application that each sensor transmits data to the BS every 25 seconds. The length of the data packet is fixed at 2048 bits and the initial energy of each node is $10^{-3}$J. The transmission range is 16 m, and the scanning angle is $\frac{\pi}{4}$. To examine the performances influenced by the percentage setting in our proposed protocol, we vary the percentage in the threshold from 0.1 to 0.4 in our simulation.
D. Evaluation Metrics

The performance of the protocol is evaluated in terms of connectivity, the number of living nodes and the packet delivery ratio.

- Connectivity (CN)

\[ CN = \frac{N_C}{N_T} \quad (6) \]

where \( N_C \) is the number of sensor nodes that involved in at least one path to the BS; \( N_T \) is the total number of sensor nodes in the network.

- The Number of Living Nodes (NLN): the nodes that still have energy in the network.

- The Packet Delivery Ratio (PDR)

\[ PDR = \frac{\sum P_{rx}}{\sum P_{tx}} \quad (7) \]

where \( \sum P_{tx} \) is the total number of transmitted packets from the sensor nodes and \( \sum P_{rx} \) is the total number of received packets by the BS.

V. Results and Analysis

Fig. 7 shows the result of connectivity versus time and it is intuitive. RRP outperforms simple-link, M-SIRLoS and DRS. The connectivity of simple-link and M-SIRLoS drops dramatically as time goes by, for both of them do not consider reconfiguration in the death of nodes. During data transmission, sensor nodes that involve in more number of paths consume energy faster because they have more chances to forward packets from their predecessors. Once they die of the depletion of energy, other nodes which have to forward packets through these dead nodes will lose the contact with the BS, resulting in the sharp and discrete decrease of connectivity. However, in DRS, nodes are isolated from each other once they get the BSI, and in RRP, the dying nodes are also separated from others. The death of a node does not affect the up-link routes of others. Thus, the connectivity of both DRS and RRP shows continuous dropping trend as time goes by.

Nodes in DRS use direct communication with the BS once they know the BSI, which is different from the RRP that includes both multi-hop and direct communication. To illustrate the energy dissipated in DRS and RRP, we consider the energy expended transmitting a single \( k \)-bit message from a GS, say \( S_t \), located a distance \( d \) from the BS (shown in Fig. 8) using direct communication approach, then we have:

\[ E_{direct} = k \times \varepsilon \times \varphi \times d^2 \quad (8) \]

The energy required for multi-hop routing is as follows:

\[ E_{multi-hop} = nE_{AX}(k, r, \varphi) + nE_{RX}(k) + E_{PX}(k) \quad (9) \]

Then the direct communication requires less energy than multi-hop communication if \( E_{direct} < E_{multi-hop} \). According to Eq. (3)-(5), we have

\[ d < \sqrt{\frac{n(y^2 + \frac{E_R}{\varepsilon \times \varphi}) + \frac{E_{CCR}}{\varepsilon \times \varphi}}{\varepsilon \times \varphi}} \quad (10) \]

Eq. (10) indicates the requirement for \( d \) and \( n \) if direct communication consumes less energy than that in multi-hop communication. In DRS, the further the node away from the BS, the faster it consumes energy. Nodes which are located around the bound of the simulation area die quickly. However, in RRP, most of bound located nodes are just few hops away from the CH, which slows down the rate of energy consumption. Besides, when a node becomes a DN, the AN in RRP tries to find another short path back to the BS by re-orientation. It orients to the BS if and only if no qualified candidate in its circle area is available. This reduces the probability of direct transmission, which may cause larger energy dissipation if Eq. (10) is not satisfied. Therefore, in view of connectivity and the number of living nodes (shown in Fig. 9), RRP is better than DRS. According to the definition of network lifetime in [14], which is the period from when the nodes begin transmitting data, till the time when the network coverage area drops to 70% or below, the network lifetime is about 45% improvement over DRS if RRP is used.

Interestingly, in Fig. 9, we observe that the decreasing trend of the number of living node indicated by simple-link and M-SIRLoS is smaller than that indicated in RRP. In simple-link and M-SIRLoS, the death of a node may cause tremendous decrease of connectivity because paths containing the dead node to the BS are broken. The dead node’s successors, which are supposed to forward their packets from the dead node and the dead node’s predecessors, may have much less number of packets to transmit in this case. It helps decelerate their energy consumption rate, prolonging their lifetime. However, RRP aims to maintain the network connectivity for packet delivery. The death of a node does not affect others’ packet delivery. Thus the number of transmission per node is larger than that in simple-link and M-SIRLoS, and the number of living node in RRP presents a larger decreasing trend.

Fig. 10 shows the results of packet delivery ratio. As expected, by reconfiguring the network, the RRP maintains fairly high packet delivered ratio. This is the result of the network connectivity maintenance. DNs are isolated from the original paths, their status does not affect the packet delivery of other nodes. However, in simple-link and M-SIRLoS, as time goes by, many packets are lost due to the lack of connectivity. In DRS, as the nodes communicate with the BS independently after getting the BSI, it also maintains high packet delivery ratio.

In our simulation, we also examine the performance affected by the percentage in the threshold through varying it from 0.1 to 0.4. The results imply that larger threshold can postpone the death of the first node. This is because larger threshold symbolizes larger remaining energy for a node to reconfigure, and the communication stress of the DN can be mitigated early. However, early reconfiguration may cause faster connectivity dropping rate after certain time, such as the performance shown in RRP-0.4 after 2700s. For one thing, if a node orients to a qualified neighbor, it may release its communication pressure while it burdens the pressure of the target node. For another thing, if a node orients to the BS directly, the proba-
bility of consuming energy faster is higher if Eq. (10) is not satisfied. To a certain extent, early reconfiguration means early burdening or early accelerating energy consumption, which can explain the faster connectivity dropping rate after certain time. The percentage in the threshold should be set properly according to network traffic load and in our simulation case, 0.2 is the optimal value.

VI. CONCLUSION

In this paper, we proposed the RRP for FSOSN to maintain network connectivity, guarantee packet delivery and prolong network lifetime during data transmission. In RRP, the data is firstly transmitted through the shortest hop path. Then the reconfiguration is initiated by the node whose residual energy has reached a threshold. Those nodes affected by this dying node are classified into two types and reconfigured according to their types. An energy model was setup to evaluate the proposed protocol. The simulation results indicate that RRP can improve the connectivity and packet delivery ratio significantly compared with simple-link and M-SIRLoS by isolating the DN from the original path and reconfiguring the affected nodes. Besides, we show that RRP can achieve about 45% improvement of network lifetime over DRS by providing the insight of energy dissipation of both direct communication and multi-hop communication. Through varying the percentage value in the threshold, we show that larger threshold can postpone the death of the first node, but also can cause faster connectivity dropping rate after certain time. The threshold should be properly set according to the network traffic load.

VII. ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2010-0029432).