

Delay- and Interference-Aware Routing for Wireless Mesh Network

Yuan Chai  and Xiao-Jun Zeng , *Member, IEEE*

Abstract—Effective routing design can significantly improve the whole network performance. In order to achieve the global best network performance, the problem of routing can be formulated as a constrained optimization problem. A delay- and interference-aware routing (DIAR) method using optimization is proposed in this article to find effective routes in a wireless mesh network. With the rapid development of wireless communication, next-generation networks urge shorter delay. DIAR aims at selecting routes with minimum end-to-end delay for several concurrent data flows. Delay is derived according to interference, bandwidth, and the probability of transmission failure. Then, the relationship between delay and the number of interfering nodes is built for the first time, which makes the estimation of delay more simple. When solving the optimization problem, an improved genetic algorithm is proposed to balance load. Besides, DIAR considers dynamic network condition caused by selecting different paths to transmit packets and evaluates the network condition while finding the solution of routing. The paths with least end-to-end delay will be finally chosen as the solution. Simulation results show that DIAR can obtain better network performance.

Index Terms—Delay, interference, optimization model, routing, wireless mesh network (WMN).

I. INTRODUCTION

DUE to the features of low cost, high robustness, and reliability, a wireless mesh network (WMN) is an essential architecture in next-generation communication network, which attracts the attention of many researchers [1], [2]. The self-construction and self-configuration peculiarities can dramatically reduce the complexity of network deployment and maintenance. As the WMN can be established flexibly with low cost, it can be used in temporary networks like communication networks deployed in an exhibition, where users are temporarily crowded. Besides, the WMN can recover communication fast when disasters or accidents happen, so it is very suitable for emergency communication, military communication, enterprise wireless communication, and so on. Smart home devices can also build a WMN, and multiple hops can extend the coverage area. Although communication technology develops very fast, and 5G is being used, there are still many rural areas, where it is too hard and costly to deploy traditional base stations. These scattered areas with complex physiognomy are far away from cities and towns so that the traditional deployment will cost a lot.

Manuscript received May 31, 2019; revised October 25, 2019; accepted January 10, 2020. (*Corresponding author: Xiao-Jun Zeng.*)

The authors are with the School of Computer Science, University of Manchester, Manchester M13 9PL, U.K. (e-mail: yuan.chai-2@postgrad.manchester.ac.uk; x.zeng@manchester.ac.uk).

Digital Object Identifier 10.1109/JSYST.2020.2966795

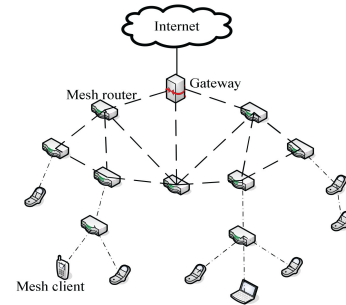


Fig. 1. Network model of the WMN.

In this case, the WMN is valuable to provide Internet service. The network model is shown in Fig. 1.

There are two types of mesh nodes in the WMN: mesh routers and mesh clients [3], [4]. Mesh routers are always equipped with multiple radio interfaces connected with multiple wireless channels, while mesh clients have a single radio interface. A common channel must be set for both mesh routers and clients to communicate between them smoothly. The data traffic in the WMN includes gateway-oriented and client-oriented traffic [5]. In gateway-oriented traffic, data flows are convergent through the gateway to the Internet. In client-oriented traffic, different mesh nodes are connected by multiple hops.

Routing is very important in network design to improve network performance [6]. Bad path selections will cause congestion, high interference, long delay, and so on. Thus, designing an effective routing method considering link quality is essential [7]. Nowadays, more and more communication services require little delay, so a delay- and interference-aware routing (DIAR) method minimizing end-to-end delay is proposed in this article. The main contributions of DIAR can be summarized as follows.

1) Derive the mathematical relationship between delay and the number of interfering nodes, and formulate routing into an optimization problem. As the delay is a crucial criterion of network performance, minimizing delay is the objective of the routing optimization model. As a result, the route with the least end-to-end delay will be selected. To evaluate delay accurately, DIAR considers bandwidth, interference, and the probability of transmission failure (PTF) at the same time. Both transmission and backoff delay are considered. As directly obtaining delay by a proactive method will cause high cost, establishing the relationship between delay and fundamental network factors is vital. As far as we know, DIAR first builds the relationship between delay and the number of interfering nodes, which can estimate delay and select paths effectively with low cost.

2) Extend the basic genetic algorithm (GA) to balance the load when solving the routing optimization problem. The basic GA finds solutions by random crossover and mutation, and it

cannot obtain the routing solution effectively. Heavy congestion and load will occur if more than one data flow use the same mesh node, and the basic GA does not consider this particular condition. DIAR overcomes this weakness, which produces a more adaptive method to solve the routing problem. If a common node is used by multiple data flows, another mesh node with the least number of interfering nodes will be selected as a replacement node. The rationality of this method is given according to the derived relationship between delay and the number of interfering nodes. In addition, as different selected routes can influence and change network condition in turn, the improved GA considers the dynamic condition as input parameters and evaluates the condition iteratively to find optimal paths.

II. RELATED WORKS

The research of routing optimization designed for the WMN includes the formulation of the optimization model and solution of the optimization model. Some research studies do not particularly consider interference. A search-based routing using GA [8] is proposed in [9]. This method considers expected transmission time (ETT) [10] in addition to hop count as the routing metric. The fitness function is based on hop count and aggregated ETT. The end-to-end delay, according to ETT, is minimized, and results are better than the traditional hop count metric. The performance of GA in the context of routing in the WMN is then evaluated. A robust GA-based quality-of-service (QoS) routing method is proposed in [11]. Multiple feasible paths are found in this method, which can improve the robustness. For each chromosome, the normalized cost based on all types of QoS metrics is used as the fitness value. Different QoS parameters can be considered at the same time, but the link metrics are generated uniformly. The routing algorithm proposed in [12] considers delay, throughput, and packet error ratio separately according to different types of communication applications. The interaction between medium access control (MAC) and routing algorithms is considered sufficiently. Besides, the resource for QoS flows is reserved. The routing algorithm proposed in [13] uses expected transmission count (ETX) [14] to evaluate link conditions and develops a multiobjective approach for the routing problem. The modified nondominated sorting GA is used to discover better and diverse routing solutions.

Some other research studies consider the interference influence. The routing method in [15] maximizes throughput by analyzing different types of interference. The two types of interference are distinguished by the distance between the interferer link and the reference link. Then, the maximum throughput that the network can support is obtained. The mixed-integer linear program model is established to formulate this routing problem. Spatial reusability-aware single-path routing (SASR) and spatial reusability-aware anypath routing (SAAR) [16] consider the spatial reusability problem in routing. The interference relationship is influenced by the node location. The path with the minimized cost, which is composed of noninterfering links, is selected. Load-balance and interference-aware (LBIA) [17] considers the load balance problem in the situation of multicast sessions, and the link with less load is selected. Intraflow and interflow interference are considered separately in LBIA. The process of finding paths in LBIA is the approach of building a multicast tree. The method proposed in [18] is based on a greedy algorithm [19], and it can choose the path with the minimum cost by considering the usage of wireless channels. An end-to-end throughput model based on the IEEE 802.11 protocol [20] is

proposed in [21], which can guide routing design. Load-aware route selection (LARS) [22] selects load-balance paths and gateways by using a queuing model and estimating the residual capacity of network paths. Network bottlenecks are identified, and the network paths to upstream Internet flows are allocated to improve the network capacity. In this way, wireless network resources and gateways are utilized in a balanced way. Distributed delay-aware routing (DDAR) [23] is a multihop multipath routing in a joint distributed scheduling and routing problem. Based on the conflict graph model and Lyapunov optimization, the delay of each commodity flow is minimized. Although these research studies consider interference, the particular way to obtain delay with low cost is not derived. The cost of obtaining network conditions is not sufficiently considered.

Some research studies typically focus on finding effective solutions for optimization models. Jiang *et al.* [24] consider delay, bandwidth, and packet loss ratio to satisfy the communication request. The solution algorithm of this article combines GA and ant colony optimization (ACO) [25]. The algorithm proposed in [26] minimizes the cost to guarantee the QoS request, and it combines GA and particle swarm optimization (PSO) [27] to solve the optimization problem. Cuckoo search optimization ad-hoc on-demand distance vector (CSO-AODV) [28] applies Cuckoo Search and uses the best fitness value computation to find the best path when multiple routes are available. It satisfies QoS constraints like load, energy, and hop count when discovering routes. However, the congestion caused by using a common node in a particular routing problem is not effectively taken into account in these research studies.

In general, existing research studies satisfy different focuses, but the approach and cost of obtaining network conditions are overlooked. Collecting information by probe packets can cause high load and cost, which will degrade the network performance. So, finding an effective and easy way to evaluate link performance is essential. Besides, when the solution of routing is different, the network condition will also be different. That is, the solution can influence the network condition in turn. The process of finding routes and evaluating network condition should be done at the same time.

III. SYSTEM MODEL

The WMN is modeled by a network connectivity graph $G = (V, L)$, where V and L are the sets of all mesh nodes and links in the WMN, respectively. Nodes are deployed uniformly, and multiple wireless channels for mesh routers are allocated randomly. The set of source-destination traffic (denoted as F) is given. For the sake of clarity, all notations used in this article are listed in Table I.

A. Transmission Model

The transmission model considers both path loss and fading channel models. The path loss model is chosen according to the distance between the source and the destination [31]. If the distance is shorter than the critical distance, the free space propagation model will be used. If the distance is larger than the critical distance, then the two-ray ground reflection model will be selected. The critical distance (denoted as d_0) is

$$d_0 = \frac{4\pi h_t h_r}{\lambda} \quad (1)$$

where h_t and h_r are heights of transmitting and receiving antennas; λ is the wavelength. The path loss effect between source

TABLE I
VARIABLE NOTATIONS

Notation	Denotation
V	Set of mesh nodes
L	Set of wireless links in WMN
F	Traffic of data flows
h_t	Height of transmitting antennas
h_r	Height of receiving antennas
λ	Wavelength
G_t	Gain of transmitting antennas
G_r	Gain of receiving antennas
d_0	Critical distance of different transmission models
i	The source node of a link
j	The destination node of a link
d_{ij}	Distance between source node i and destination node j
PL_{ij}	Path loss between i and j
$E(p)$	Channel fading between i and j
P_{ij}^{loss}	Receiving power only considering path loss
P_{ij}^{fading}	Receiving power only considering channel fading
(i,j)	A wireless link
σ_{ij}^2	Power of multipath components on link (i,j)
e	A data flow
X_{ij}^e	(binary)=1, when link (i,j) is used to service flow e
N_i^e	The number of queuing packets at node i in data flow e
$T_{(i,j)}^e$	Delay cost by transmitting each packet on link (i,j) for flow e
S	Set of source nodes
D	Set of destination nodes
c_{ij}^e	(binary)=1, when link (i,j) exists
$P_{(i,j)}$	Probability of transmission failure on link (i,j)
$Length$	Size of data packet
K	Largest allowed number of retransmission
k	Total number of transmission
$B_{(i,j)}$	Available bandwidth of link (i,j)
$SlotTime$	Time of a slot
CW_s	Duration of the s th backoff window
CW_{min}	Minimum backoff window
P_{T_i}	Transmission power of source node i
P_{ij}	Receiving power at destination node j from source node i
P_{a_j}	Receiving power of interference at node j from node a
I_{ij}	Interference to link (i,j)
γ_{ij}	SINR of link (i,j)
N	Power of background noise
B_0	Nominal bandwidth
r_T	Transmission range
r_I	Interference range
β	Threshold of SINR
I_j	Set of interfering nodes of node j
n	The number of interfering nodes of receiving node j
δ	Adjustment parameter
ξ	Difference between critical values of piecewise function

node i and destination node j (denoted PL_{ij}) can be expressed as

$$PL_{ij} = \frac{P_{ij}^{loss}}{P_{T_i}} = \begin{cases} \frac{G_t G_r \lambda^2}{(4\pi d_{ij})^2}, & 0 < d_{ij} \leq d_0 \\ \frac{G_t G_r h_t^2 h_r^2}{(d_{ij})^4}, & d_0 < d_{ij} \leq r_T \end{cases} \quad (2)$$

where P_{ij}^{loss} is the receiving power at destination node j from source node i , which is affected by path loss. P_{T_i} is the transmission power of source node i . G_t and G_r are the gain of transmitting and receiving antennas, respectively. d_{ij} is the distance between source node i and destination node j .

To describe the realistic channel feature, the Rayleigh fading channel model [32] is also used. The envelope of the received signal from node i to node j (denoted as x_{ij}) is Rayleigh distributed, and the probability density function of x_{ij} (denoted as $f_X(x_{ij})$) is

$$f_X(x_{ij}) = \frac{x_{ij}}{\sigma_{ij}^2} e^{-\frac{x_{ij}^2}{2\sigma_{ij}^2}}, \quad x_{ij} \geq 0 \quad (3)$$

where σ_{ij}^2 is the power of multipath components.

Let received signal power $p_{ij} = x_{ij}^2$, so the distribution of p_{ij} (denoted as $f_P(p_{ij})$) is

$$f_P(p_{ij}) = \frac{dx_{ij}}{dp_{ij}} f_X(\sqrt{p_{ij}}) = \frac{1}{2\sigma_{ij}^2} e^{-\frac{p_{ij}}{2\sigma_{ij}^2}}, \quad p_{ij} \geq 0. \quad (4)$$

Therefore, the received signal power is exponent distributed, and its expected value (denoted as $E(p_{ij})$) is

$$E(p_{ij}) = 2\sigma_{ij}^2. \quad (5)$$

Based on (5), the average receiving power only influenced by channel fading (denoted as P_{ij}^{fading}) is

$$P_{ij}^{fading} = P_{T_i} \cdot E(p_{ij}) \quad (6)$$

where P_{T_i} has the same meaning as that in (2).

According to the path loss and channel fading effects, the final received signal power (denoted as P_{ij}) can be expressed as

$$P_{ij} = P_{T_i} \cdot PL_{ij} \cdot E(p_{ij}). \quad (7)$$

B. MAC Protocol

The IEEE 802.11 standard is used as the MAC-layer protocol, and the distributed coordination function is applied [20]. To avoid collisions of the contending wireless channel, the backoff mechanism is used. If a packet is not successfully received by the destination, the source will retransmit until the number of retransmission reaches the maximum value. Before sending packets, the node detects channel condition lasting a random number of backoff window. The maximum range of the backoff window will be doubled in each retransmission. When the times of retransmission reaches the maximum value, the packet will be dropped.

IV. DIAR MODEL DESIGN

DIAR is a type of centralized routing method, where the gateway and network manager makes the selection of routes according to the statistics and topology information collected from each mesh node. The network manager will then do the calculation and send back the best route to mesh nodes [22], [29]. DIAR formulates the route selection issue as an optimization problem and overcomes the current weaknesses. Signal-to-interference-plus-noise ratio (SINR) [30], bandwidth, PTF, and delay can describe the network condition, and delay is influenced by the other three network factors. To overcome the difficulty of obtaining delay, we derive the mathematical expression of PTF according to the number of interfering nodes. Then, as the delay is related to PTF, the relationship between delay and the number of interfering nodes is first given in DIAR. Besides, DIAR improves the GA to balance load in the typical routing problem. When different paths and combination of wireless links are selected, the number of interfering nodes for each used node will be different, and the network condition and delay will change. The improved GA method also evaluates the dynamic network condition in each iteration. The optimization process and the GA are started every time when a flow demand arrives. When a new traffic flow arrives, the source node will send a request message to the network manager, which makes the new flow be detected. The network manager will then execute the optimization process. Finally, DIAR can select the effective routes with less delay, less interference, less PTF, and larger available bandwidth for multiple data flows at the same time, which can achieve better whole network performance.

A. Establishment of the Optimization Model

1) *Objective*: Minimizing end-to-end delay is the objective of the optimization model of DIAR, and it is given as

$$\text{minimize } \left\{ \sum_{(i,j) \in L} \sum_{e \in F} X_{ij}^e T_{(i,j)}^e (N_i^e + 1) \right\} \quad \forall e \in F \quad (8)$$

where L is the set of links in the network, e is a data flow, and F is the set of all data flows. X_{ij}^e is the binary decision variable of the optimization problem to show whether flow e is via link (i, j) . If link (i, j) is chosen to transmit packets of data flow e , X_{ij}^e is 1, or X_{ij}^e is 0. $T_{(i,j)}^e$ is the delay cost by transmitting each packet on link (i, j) for data flow e . N_i^e is the number of queuing data packets waiting to be transmitted at node i in data flow e . For each data flow, the end-to-end delay will be minimized, and it indicates the goodness level of routing solution. The detailed mathematical expression of delay will be given in Section V.

2) *Constraints*: Constraints will guarantee smooth communication. The detailed constraints of DIAR are listed as follows.

The source node of each data flow must connect one neighbor node to transmit data packets out

$$\sum_{(i,j) \in L} X_{ij}^e = 1, \quad i \in S, \forall e \in F \quad (9)$$

where S is the set of source nodes. Similarly, the destination node of each data flow also must connect one neighbor node to receive data packets

$$\sum_{(i,j) \in L} X_{ij}^e = 1, \quad j \in D, \forall e \in F \quad (10)$$

where D is the set of destination nodes. Then, the intermediate nodes should guarantee that all receiving packets can leave through another link. This request is shown as

$$\sum_{(i,j) \in L} X_{ij}^e = \sum_{(j,u) \in L} X_{ju}^e \quad \forall j \in V - \{S, D\}, \forall e \in F \quad (11)$$

where u is the next hop of node j in data flow e . Only one node can be selected as the next hop. Formulas (9)–(11) represent the flow conservation. If a loop exists in a route, there will be an intermediate node with different number of input and output links, that is, $\sum_{(i,j) \in L} X_{ij}^e \neq \sum_{(j,u) \in L} X_{ju}^e$, which is conflicted with the constraint (11). Thus, the path identified by DIAR, which meets constraints (9)–(11) is loop free.

Besides, the wireless links selected to be used must exist in the WMN

$$c_{ij}^e \geq X_{ij}^e \quad \forall (i, j) \in L, \forall e \in F \quad (12)$$

where c_{ij}^e is a binary constant to show whether link (i, j) exists in the network. If the distance between nodes i and j is within the transmission range of each other, then link (i, j) exists and the value of c_{ij}^e is 1. Otherwise, link (i, j) cannot appear, and the value of c_{ij}^e is 0. This constraint means that the end nodes of selected links must be within the transmission range of each other.

The value of variable X_{ij}^e in the optimization model is binary

$$X_{ij}^e \in \{0, 1\} \quad \forall (i, j) \in L, \forall e \in F. \quad (13)$$

Then, the optimization model is formulated with the objective given in (8), subject to the constraints given (9)–(13). Based on smooth communication, the end-to-end delay of the whole network is minimized during the process of routing.

B. Solution of the Optimization Model

Due to the interaction between the solution of routing and the network condition, the optimization problem formulated in the last two subsections is difficult to be solved by the conventional optimization algorithms such as mathematical programming. On the other hand, the GA is applicable to the given optimization problem, and it has been successfully applied to solve some routing problems, as stated in Section II. The GA is a very effective approach to find an optimal solution [33], and it is even regarded as the best choice to find the optimal path [9]. For these reasons, the GA is used here to solve the formulated optimization model. The GA obtains the solution by iterations, and it will stop when either the maximum number of iterations is reached or no improvement can be achieved by further iterations. The set of solutions in each iteration is known as a population. Each population includes some chromosomes, and every chromosome represents the selected path for data flows. The chromosome is composed of a sequence of node numbers, which are the ones selected to transmit data packets. Some random and valid paths are set as the initial population. As DIAR considers the dynamic network condition, the delay of selected links is evaluated and updated for each chromosome in the iterative optimization.

Crossover and mutation can explore new solutions. In DIAR, to avoid some elite parent chromosomes becoming worse after crossover and mutation, all parent and offspring chromosomes are sorted together according to their levels of goodness. The worst chromosomes that are over the population size will be deleted. Besides, when duplicated chromosomes are produced, only one will be kept; the others will be deleted.

The quality and goodness of chromosomes in the constrained optimization problem is evaluated by the fitness function (denoted as $\text{Fit}(i)$), which is

$$\text{Fit}(i) = \begin{cases} \frac{T(i) - T_{\min}}{T_{\max} - T_{\min}}, & J_{\text{voi}}(i) = 0 \\ \sqrt{\frac{\sum_{j=1}^J \left(\frac{\max(0, f(j))}{G_{\max}^j} \right)^2}{J_{\text{voi}}(i)}}, & J_{\text{voi}}(i) \neq 0 \end{cases} \quad (14)$$

where $J_{\text{voi}}(i)$ is the number of violative constraints in chromosome i . $T(i)$ is the end-to-end delay value of chromosome i ; therefore, it is the objective of the optimization model shown in (8). T_{\min} and T_{\max} are the minimum and maximum delay values among all chromosomes of the generation. $f(j)$ is the degree of violation (absolute value of the difference between both equation sides) in terms of constraint j . If the j th constraint is satisfied, $f(j)$ is 0. G_{\max}^j is the maximum violation value of constraint j in the generation. J is the total number of constraints in the optimization model.

If $J_{\text{voi}}(i)$ is 0, the chromosome is feasible. The constraints from (9) to (13) are satisfied, where flow conservation and transmission rule are met. The fitness function can show the goodness of chromosomes. The chromosome with less fitness value is better, as it means that the corresponding feasible solution has the smaller delay. If $J_{\text{voi}}(i)$ is not 0, the chromosome is nonfeasible. The fitness function is used to show the degree of constraint violation. The smaller value of the fitness function means that a nonfeasible solution has less violation degree. When sorting the chromosomes, the reasonable chromosomes are followed by the nonreasonable ones. Each type of chromosomes is sorted by the fitness function values from small to large. The first chromosome will be considered as the best solution.

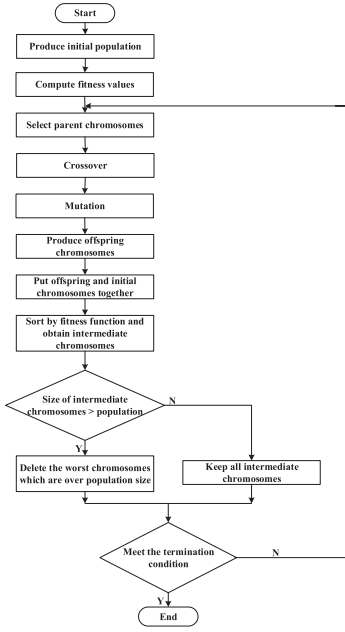


Fig. 2. Flowchart of the proposed GA.

This process of finding better solutions is executing until the maximum number of iterations is reached. Applying elitism also helps to make solutions better and better. Although the final verdict of the least delay cannot be given in advance to stop the iterations, optimal solutions usually can be obtained at the current state of the art of computation [9]. Finally, the reasonable solution with minimum fitness value (i.e., delay) will be selected as the optimal path.

The original GA finds a solution by random crossover and mutation, which ignores some particular features of routing in WMN. Such negligence could lead to some undesired performance. For example, a high congested route, which uses the same node by multiple data flows, can be obtained by the original GA. To avoid this, DIAR improves the GA to solve the optimization model. The main improvements are as follows.

1) *Crossover*: Single-point crossover [39] is used in DIAR. To ensure that the paths generated by crossover are still valid paths, the selected pair of parent chromosomes must have at least one common node in addition to the source and destination nodes. The common node of two parent chromosomes is then used as the crossover point. The former and latter parts of each chromosome are exchanged to build two new chromosomes. For instance, if the chromosomes [1-3-4-8-0] and [1-2-4-7-9-0] are selected as parent chromosomes, common node 4 will be the crossover point. [1-3-4-7-9-0] and [1-2-4-8-0] are obtained after crossover.

2) *Mutation*: In original mutation, one random point in the selected chromosome is replaced with another random neighbor node, which does not particularly consider the congestion situation in the routing problem. However, when a common node is used by multiple data flows, the queuing packets at this common node will increase, and congestion will be caused. Thus, the basic GA needs to be improved to find routes more effectively. Based on the relationship between delay and the number of interfering nodes, which will be given in detail in Section V, when one node is used by more than one data flows, the improved algorithm in DIAR will choose another neighbor node with least number of interfering nodes as a replacement node. For example,

if five flows (flows 1-5) are transmitted to the next node A , one of them (e.g., flow 1) will be still transmitted to A , and the other four (flows 2-5) will try to find other alternative nodes to replace the previous next node A . Ideally, the four alternative nodes chosen by flows 2-5 are different. However, if there are still some flows selecting the same alternative node B (e.g., flows 2-5 select node B at the same time in the worst-case scenario), one flow will transmit to the alternative node B (e.g., flow 2), and others (flows 3-5) will try to find another new replacement node again. This process will be executed until all flows have new and different alternative nodes or in imperfect case, until obtaining best performance improvement. The imperfect case is that all allowed alternative nodes have already been used when a flow looks for an alternative node, so a shared node with other flows has to be selected. In this case, the shared node with the least number of interfering nodes will be chosen. The best performance improvement is then obtained in the imperfect conditions. Congestion will be avoided, and load can be balanced effectively in this way. To guarantee the normal and smooth communication, the new replacement node is within the transmission range of both previous and next hops of the common node.

V. KEY PARAMETERS AND RELATIONSHIPS IN THE OPTIMIZATION MODEL OF DIAR

A. Expression of Delay

Transmission and backoff time compose the delay of transmitting each data packet successfully on the link (i, j) (denoted as $T_{(i,j)}$) [34]. The transmission time is the period cost by emptying a packet, and the backoff time is the waiting time before transmission. Therefore, the expected value of delay (denoted as $E[T_{(i,j)}]$) can be expressed as

$$\begin{aligned}
 E[T_{(i,j)}] &= \sum_{k=1}^{K+1} p_{(i,j)}^{k-1} (1 - p_{(i,j)})^{I\{k < K+1\}} \\
 &\quad \cdot \sum_{s=1}^k \left(E[CW_s] \cdot \text{SlotTime} + \frac{\text{Length}}{B_{(i,j)}} \right) \\
 &= \frac{\text{Length}}{B_{(i,j)}} \left[\frac{1 - p_{(i,j)}^K}{1 - p_{(i,j)}} \right] + E[\text{backoff}] \quad (15)
 \end{aligned}$$

where $p_{(i,j)}$ is the PTF on the link (i, j) , Length is the size of a data packet, K is the largest allowed number of retransmission, k is the total number of transmission, and $B_{(i,j)}$ is the bandwidth of link (i, j) . The indicator $I(A)$ is equal to 1 if A is true. When the number of retransmission is beyond the maximum allowed number K , the packet will be dropped. SlotTime is the time of a slot. CW_s is the size of the s th backoff window.

The packet may be transmitted successfully in the k th transmission, and the probability of the k th transmission being successful is $p_{(i,j)}^{k-1} (1 - p_{(i,j)})^{I\{k < K+1\}}$. Then, the expected delay value of a successful transmission is the sum of all cases [the successful transmission may happen in the first to $(K+1)$ th attempt].

CW_s can be expressed as

$$CW_s = 2^{s-1} \cdot CW_{\min} \quad (16)$$

where CW_{\min} is the minimum backoff window. Then, the expected value of CW_s (denoted as $E[CW_s]$) is

$$E[CW_s] = \frac{CW_s - 1}{2} = \frac{2^{s-1} \cdot CW_{\min} - 1}{2}. \quad (17)$$

So, the expected value of backoff time (denoted as $E[\text{backoff}]$) is

$$\begin{aligned} E[\text{backoff}] &= \sum_{k=1}^{K+1} p_{(i,j)}^{k-1} (1 - p_{(i,j)})^{I\{k < K+1\}} \sum_{s=1}^k E[CW_s] \\ &= \frac{CW_{\min}[1 - (2p_{(i,j)})^{K+1}]}{2(1 - 2p_{(i,j)})} - \frac{1 - p_{(i,j)}^K}{2(1 - p_{(i,j)})}. \end{aligned} \quad (18)$$

Thus, according to (15) and (18), the expected value of delay of link (i, j) (denoted as $E[T_{(i,j)}]$) can be expressed as

$$\begin{aligned} E[T_{(i,j)}] &= \frac{\text{Length}}{B_{(i,j)}} \left[\frac{1 - p_{(i,j)}^K}{1 - p_{(i,j)}} \right] \\ &+ \frac{CW_{\min}[1 - (2p_{(i,j)})^{K+1}] \cdot \text{SlotTime}}{2(1 - 2p_{(i,j)})} \\ &- \frac{(1 - p_{(i,j)}^K) \cdot \text{SlotTime}}{2(1 - p_{(i,j)})}. \end{aligned} \quad (19)$$

Delay of link (i, j) is, therefore, influenced by the available bandwidth (i.e., $B_{(i,j)}$) and the PTF (i.e., $p_{(i,j)}$). The approach of obtaining these two factors in DIAR is shown as follows.

B. Available Bandwidth

According to (7), the receiving power at destination node j from source node i (denoted as P_{ij}) is related to the path loss and channel fading between i and j , and it is

$$P_{ij} = \begin{cases} 2P_{Ti} \cdot \sigma_{ij}^2 \cdot \frac{G_t G_r \lambda^2}{(4\pi d_{ij})^2}, & 0 < d_{ij} \leq d_0 \\ 2P_{Ti} \cdot \sigma_{ij}^2 \cdot \frac{G_t G_r h_t^2 h_r^2}{(d_{ij})^4}, & d_0 < d_{ij} \leq r_T \end{cases} \quad (20)$$

where P_{Ti} is the transmission power of source node i .

The total interference to link (i, j) (denoted as I_{ij}) is

$$I_{ij} = \sum_{\forall a \in V, a \neq i} \sum_{\forall b \in V, b \neq j} P_{Ta} \cdot PL_{aj} \cdot E(p_{aj}) \cdot X_{ab} \cdot X_{ij} \quad \forall i, j \in V, i \neq j \quad (21)$$

where a and b are the source and destination nodes of interfering link (a, b) , respectively. Interference exists when interfering links and data transmission link (i, j) are used at the same time. The interference to link (i, j) is the sum of interference caused by all interfering links. According to the receiving power and interference, the SINR of link (i, j) (denoted as γ_{ij}) is

$$\gamma_{ij} = \frac{P_{ij}}{N + I_{ij}} \quad \forall i, j \in V, i \neq j \quad (22)$$

where N is the power of background noise in the network. Then, the available bandwidth (denoted as $B_{(i,j)}$) can be expressed as

$$B_{(i,j)} = B_0 \cdot \log_2(1 + \gamma_{ij}) \quad \forall i, j \in V, i \neq j \quad (23)$$

where B_0 is the nominal bandwidth.

C. PTF

When the SINR is lower than the threshold, the transmission will be failed. The threshold is the minimum requested value of the SINR that can guarantee data packets to be decoded successfully by the receiving node. PTF is, thus, the probability of the SINR lower than the threshold. As the SINR is influenced by the receiving data signal power (i.e., P_{ij}) and interference (i.e., P_{aj}) based on (22), PTF is also related to both P_{ij} and P_{aj} .

1) *Probability Density Function of Receiving Power P_{ij}* : As P_{ij} is related to the distance between transmitting node i and receiving node j (i.e., d_{ij}), the cumulative distribution function of P_{ij} can be derived according to the cumulative distribution function of d_{ij} . The cumulative distribution function of d_{ij} is

$$p(d_{ij} < r_i) = \frac{\pi r_i^2}{\pi r_T^2} = \frac{r_i^2}{r_T^2}, \quad 0 < r_i \leq r_T \quad (24)$$

where r_i is a particular distance from source node i , and r_T is the transmission range. The probability density function of d_{ij} is

$$f_{d_{ij}}(r_i) = \frac{2r_i}{r_T^2}, \quad 0 < r_i \leq r_T. \quad (25)$$

Based on the relationship between P_{ij} and d_{ij} in (20) and the cumulative distribution function of d_{ij} in (24), the cumulative distribution function of P_{ij} can be expressed as

$$p(P_{ij} < P_x) = \begin{cases} p\left(d_{ij} > \sqrt{\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{16\pi^2 P_x}}\right) \\ = 1 - \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{16\pi^2 r_T^2 P_x}, & \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{16\pi^2 d_0^2} \leq P_x < +\infty \\ p\left(d_{ij} > \sqrt[4]{\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{P_x}}\right) \\ = 1 - \sqrt{\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{\frac{P_x}{r_T^2}}}, & \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{r_T^4} \leq P_x < \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{d_0^4}. \end{cases} \quad (26)$$

Thus, the probability density function of P_{ij} can be derived as

$$f_{P_{ij}}(P_x) = \begin{cases} \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{16\pi^2 r_T^2} \cdot P_x^{-2}, & \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{16\pi^2 d_0^2} \leq P_x < +\infty \\ \frac{\sqrt{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}}{2r_T^2} \cdot P_x^{-\frac{3}{2}}, & \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{r_T^4} \leq P_x < \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{d_0^4}. \end{cases} \quad (27)$$

2) *Probability Density Function of Interference Power P_{aj}* : P_{aj} is the power of interference at node j from node a . The process of getting the probability density function of P_{aj} is similar to P_{ij} .

The cumulative distribution function of distance d_{aj} is

$$p(d_{aj} < r_a) = \frac{\pi r_a^2}{\pi r_I^2} = \frac{r_a^2}{r_I^2}, \quad 0 < r_a \leq r_I \quad (28)$$

where r_a is a particular distance from interfering source node a , and r_I is the interference range. Similar to the process of obtaining cumulative distribution function of P_{ij} , the cumulative distribution function of P_{aj} is

$$p(P_{aj} < P_a) = \begin{cases} 1 - \frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 r_I^2 P_a}, \frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 d_0^2} \leq P_a < +\infty \\ 1 - \sqrt{\frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{\frac{P_a}{r_I^2}}}, & \frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{r_I^4} \leq P_a < \frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{d_0^4}. \end{cases} \quad (29)$$

The probability density function of P_{aj} is then derived as

$$f_{P_{aj}}(P_a) = \begin{cases} \frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 r_I^2} \cdot P_a^{-2}, & \frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 d_0^2} \leq P_a < +\infty \\ \frac{\sqrt{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}}{2r_I^2} \cdot P_a^{-\frac{3}{2}}, & \\ \frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{r_T^4} \leq P_a < \frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{d_0^4}. & \end{cases} \quad (30)$$

3) *Expression of the PTF*: According to the probability density function of P_{ij} and P_{aj} shown as (27) and (30), the PTF on link (i, j) (denoted as $p_{(i,j)}$) is

$$\begin{aligned} p_{(i,j)} &= p(\gamma_{ij} < \beta) = p\left(\frac{P_{ij}}{N + \sum_{a \in I_j} P_{aj}} < \beta\right) \\ &\approx p\left(P_{ij} < \beta \sum_{a \in I_j} P_{aj}\right) = p(P_x < \beta n P_a) \\ &= \int_{-\infty}^{+\infty} f_{P_{aj}}(P_a) \cdot p(P_x < n\beta P_a | P_a) dP_a \\ &= \int_{-\infty}^{+\infty} f_{P_{aj}}(P_a) \int_{-\infty}^{n\beta P_a} f_{P_{ij}}(P_x) dP_x dP_a \quad (31) \end{aligned}$$

where β is the threshold of the SINR. I_j is the set of interfering nodes of node j . n is the number of interfering nodes of receiving node j . In the interference limited network, interference from other transmitters is much larger than the white noise at receivers, so the background noise can be ignored [35] [36].

If $\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{r_T^4} \leq n\beta P_a \leq \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{d_0^4}$, (31) can be expressed as

$$\begin{aligned} p_{(i,j)} &= \int_{-\infty}^{+\infty} f_{P_{aj}}(P_a) \cdot \left(1 - \frac{\sqrt{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}}{\sqrt{n\beta r_T^2}} P_a^{-\frac{1}{2}}\right) dP_a \\ &= \int_{\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{n\beta d_0^4}}^{\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{n\beta r_T^4}} \frac{\sqrt{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}}{2r_I^2} P_a^{-\frac{3}{2}} \\ &\quad \cdot \left(1 - \frac{\sqrt{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}}{\sqrt{n\beta r_T^2}} P_a^{-\frac{1}{2}}\right) dP_a \\ &= \frac{\sqrt{n\beta P_{Ti} P_{Ta}} (r_T^4 - 2r_T^2 d_0^2 + d_0^4) \sigma_{aj}}{2\sigma_{ij} P_{Ti} r_T^2 r_T^2}. \quad (32) \end{aligned}$$

If $\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{d_0^4} < n\beta P_a < +\infty$, (31) can be expressed as

$$\begin{aligned} p_{(i,j)} &= \int_{-\infty}^{+\infty} f_{P_{aj}}(P_a) \cdot \left(1 - \frac{\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{8\pi^2 n\beta P_a r_T^2}\right) dP_a \\ &= \int_{\frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 d_0^2}}^{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2} \frac{\sigma_{aj}^2 P_{Ta} G_t G_r \lambda^2}{8\pi^2 r_I^2} P_a^{-2} \\ &\quad \left(1 - \frac{\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{8\pi^2 n\beta P_a r_T^2}\right) dP_a \\ &\quad + \int_{\frac{\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{8\pi^2 n\beta d_0^2}}^{\frac{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}{d_0^4}} \frac{\sqrt{2\sigma_{aj}^2 P_{Ta} G_t G_r h_t^2 h_r^2}}{2r_I^2} P_a^{-\frac{3}{2}}. \end{aligned}$$

$$\begin{aligned} &\left(1 - \frac{\sigma_{ij}^2 P_{Ti} G_t G_r \lambda^2}{8\pi^2 n\beta P_a r_T^2}\right) dP_a \\ &= \frac{(3d_0^2 r_T^2 - d_0^4) \sigma_{aj} \sqrt{n\beta P_{Ta}}}{3\sigma_{ij} r_I^2 r_T^2 \sqrt{P_{Ti}}} - \frac{P_{Ti} \sigma_{ij}^2 d_0^4}{6n\beta r_I^2 r_T^2 \sigma_{aj}^2 P_{Ta}}. \quad (33) \end{aligned}$$

Thus, the relationship between the PTF (i.e., $p_{(i,j)}$) and the number of interfering nodes (i.e., n) is built. As the PTF can influence the delay, which is shown in (19), the relationship between delay and the number of interfering nodes is, therefore, established. Although different nodes have different channel conditions, the power gain of channel fading is relatively stable. Many pieces of research typically assume that all links have a unit power gain of Rayleigh fading [37], [38]. Therefore, after network configuration, the number of interfering nodes is an essential factor affecting PTF. DIAR can then evaluate delay by using the number of interfering nodes, which makes the process of obtaining delay easier. According to the position of mesh nodes and interference range, the number of interfering nodes can be obtained. Thus, DIAR can obtain delay easily without extra overhead caused by probe packets as usual.

DIAR takes the dynamic condition into account and checks the number of interfering nodes every time of finding optimal paths in the iterative optimization. Finally, paths for different data flows are selected to improve network performance globally.

VI. REASONABILITY ANALYSIS OF DIAR

The reasonability of DIAR, including the explanation of the improved solution method and the cost of DIAR, is given in this section.

A. Relationship Between the PTF and the Number of Interfering Nodes

According to (32), the derivative of $p_{(i,j)}$ for n is

$$p_{(i,j)}' = \frac{\sqrt{\beta P_{Ti} P_{Ta}} (r_T^2 - d_0^2)^2}{4P_{Ti} r_I^2 r_T^2} n^{-\frac{1}{2}}. \quad (34)$$

As n is the number of interfering nodes, $n > 0$. Besides, based on the physical significance of each parameter, all of them are greater than 0. Thus, $p_{(i,j)}' > 0$, and $p_{(i,j)}$ is a monotone increasing function of n .

In addition, according to (33), the derivative of $p_{(i,j)}$ for n is

$$\begin{aligned} p_{(i,j)}' &= \frac{2\pi h_t h_r \sigma_{aj} d_0 \sqrt{P_{Ta}} (3r_T^2 - d_0^2)}{3\lambda r_I^2 \sigma_{ij} \sqrt{\beta P_{Ti} r_T^2}} n^{-\frac{1}{2}} \\ &\quad + \frac{\sigma_{ij}^2 P_{Ti} d_0^4}{6r_I^2 r_T^2 \beta \sigma_{aj}^2 P_{Ta}} n^{-2}. \quad (35) \end{aligned}$$

Similarly, as n and other parameters are greater than 0 and r_T is greater than d_0 , $p_{(i,j)}' > 0$. Thus, $p_{(i,j)}$ is also a monotone increasing function of n in this case. Both (32) and (33) are monotone increasing functions of n . The relationship between the PTF and the number of interfering nodes is a piecewise function combining (32) and (33). Although each one has monotonicity, such a property may not hold after the combination. Therefore, the monotonicity is checked after the combination of both cases.

If n is a real number, the critical value of n is $\frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{\beta P_a d_0^4}$. However, as n is the number of interfering nodes, it is a positive integer. Then, we use an adjustment parameter δ ($0 \leq \delta < 1$) to

obtain the critical value of n in different cases. For any δ , the maximum value of (32) (denoted as $p_{1\max}$) is

$$p_{1\max} = p_1 \left(n = \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{\beta P_a d_0^4} - (1 - \delta) \right) \\ = \frac{(r_T^2 - d_0^2)^2 \sigma_{aj} \sqrt{16\sigma_{ij}^2 P_{Ti}^2 G_t G_r h_t^2 h_r^2 P_{Ta} - 8d_0^4 \beta P_a P_{Ti} P_{Ta} (1 - \delta)}}{4\sqrt{2}\sigma_{ij} P_{Ti} r_T^2 d_0^2 \sqrt{P_a}}. \quad (36)$$

The minimum value of (33) (denoted as $p_{2\min}$) can be expressed as

$$p_{2\min} = p_2 \left(n = \frac{2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2}{\beta P_a d_0^4} + \delta \right) \\ = \frac{(3d_0^2 r_T^2 - d_0^4) \sigma_{aj} \sqrt{P_{Ta} (2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2 + \delta d_0^4 \beta P_a)}}{3\sigma_{ij} r_T^2 d_0^2 \sqrt{P_{Ti} P_a}}. \quad (37)$$

So, the difference between $p_{2\min}$ and $p_{1\max}$ (denoted as ξ) is

$$\xi = p_{2\min} - p_{1\max} \\ = \frac{2d_0^2 (r_T^2 - d_0^2) \sigma_{aj} \sqrt{P_{Ti} P_{Ta} (2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2 + \delta d_0^4 \beta P_a)}}{6\sigma_{ij} P_{Ti} r_T^2 d_0^2 \sqrt{P_a}} \\ - \frac{3(r_T^2 - d_0^2)^2 \sigma_{aj} \sqrt{P_{Ti} P_{Ta} (2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2 + \delta d_0^4 \beta P_a - d_0^4 \beta P_a)}}{6\sigma_{ij} P_{Ti} r_T^2 d_0^2 \sqrt{P_a}}. \quad (38)$$

To simplify (38), let $A = P_{Ti} P_{Ta} (2\sigma_{ij}^2 P_{Ti} G_t G_r h_t^2 h_r^2 + \delta d_0^4 \beta P_a)$. According to the values of each parameter in A , $A > 0$.

Then, ξ can be expressed as

$$\xi = \frac{\sigma_{aj} (r_T^2 - d_0^2) \left(2d_0^2 \sqrt{A} - 3(r_T^2 - d_0^2) \sqrt{A - P_{Ti} P_{Ta} d_0^4 \beta P_a} \right)}{6\sigma_{ij} P_{Ti} r_T^2 d_0^2 \sqrt{P_a}}. \quad (39)$$

Let $B = 2d_0^2 \sqrt{A} - 3(r_T^2 - d_0^2) \sqrt{A - P_{Ti} P_{Ta} d_0^4 \beta P_a}$; ξ can then be further simplified as

$$\xi = \frac{\sigma_{aj} (r_T^2 - d_0^2) B}{6\sigma_{ij} P_{Ti} r_T^2 d_0^2 \sqrt{P_a}}. \quad (40)$$

Since $B > 2d_0^2 \sqrt{A} - 3(r_T^2 - d_0^2) \sqrt{A} = \sqrt{A} (5d_0^2 - 3r_T^2)$, when network configuration satisfies $r_T^2 < \frac{5}{3}d_0^2$, $B > 0$. In addition, as $r_T > d_0$, $\xi > 0$, which means $p_{2\min} > p_{1\max}$. Thus, after combining (32) and (33), $p_{(i,j)}$ is still a monotone increasing function of n . In general, the PTF will increase with an increasing number of interfering nodes.

B. Relationship Between the Delay and the PTF

According to (19), the derivative of delay for $p_{(i,j)}$ is

$$E[T_{(i,j)}]' = \frac{\left((K-1)p_{(i,j)}^K - Kp_{(i,j)}^{K-1} + 1 \right) (2\text{Length} - B_{(i,j)} \text{SlotTime})}{2B_{(i,j)} (p_{(i,j)} - 1)^2} \\ + \frac{CW_{\min} \text{SlotTime} \left(2^{K+1} K p_{(i,j)}^{K+1} - 2^K (K+1) p_{(i,j)}^K + 1 \right)}{(1 - 2p_{(i,j)})^2}. \quad (41)$$

As $p_{(i,j)}$ is the PTF, $0 < p_{(i,j)} < 1$. Therefore, $E[T_{(i,j)}]' > 0$, and $E[T_{(i,j)}]$ is a monotone increasing function of $p_{(i,j)}$. That is, the delay will increase when the PTF increases. At the same

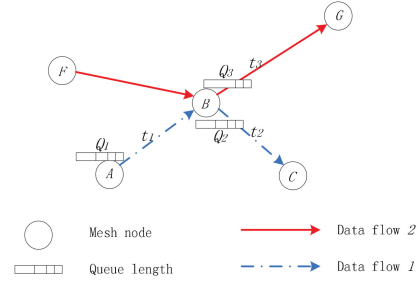


Fig. 3. Explanation of congestion and high delay at a common node.

time, as proved before, the PTF will increase with an increasing number of interfering nodes, so the delay will increase when the number of interfering nodes increases.

C. Delay of Adjacent Two Links Connected by a Common Node

An example of two data flows using a common node is shown in Fig. 3. Congestion will occur at the common node B , and the delay of serving the data flow 1 (denoted as $\text{Delay}_{\text{original}}$) is

$$\text{Delay}_{\text{original}} = t_1 + t_2 + t_3 \quad (42)$$

where t_1 is the time cost by transmitting Q_1 packets of data flow 1 from the previous hop A to the common node B . t_2 is the time cost by transmitting Q_2 packets in data flow 1 from the common node B to the next node C . t_3 is the time cost by transmitting Q_3 packets in data flow 2 from the common node B to the other next node G . Q_2 is the queue length at the common node B , which needs to be transmitted in data flow 1 to the next hop C . Q_3 is the queue length also at the common node B , which needs to be transmitted in data flow 2 to the other next hop G . The congestion occurs at node B because besides data packets of flow 1, the packets of flow 2 at the common node also need to be emptied.

To avoid the congestion and balance load, DIAR will choose a new node with the least number of interfering nodes to replace the original common node. After using another replacement node to serve data flow 1, there will be no data packets of data flow 1 waiting at the original common node B , and the situation that two data flows use the same node can be avoided. Because $p_{(i,j)}$ is a monotone increasing function of n , after choosing the node with the least number of interfering nodes as the replacement node, the new replacement link will have the least PTF. This PTF (denoted as $p_{1\text{new}}$) is

$$p_{1\text{new}} = p_1^{\min} < p_1. \quad (43)$$

Then, the new replacement link will have the least delay, and it (denoted as $t_{1\text{new}}$) is

$$t_{1\text{new}} = t_1^{\min} < t_1. \quad (44)$$

Therefore, after using the node with the least number of interfering nodes as the replacement node, the new delay (denoted as $\text{Delay}_{\text{new}}$) is

$$\text{Delay}_{\text{new}} = \text{Delay}_{\min} = t_{1\text{new}} + t_2 = t_{\min}. \quad (45)$$

The difference between the new and original delay is

$$\text{Delay}_{\text{new}} - \text{Delay}_{\text{original}} = \text{Delay}_{\min} - \text{Delay}_{\text{original}} \\ = (t_{1\text{new}} + t_2) - (t_1 + t_2 + t_3) = t_{1\text{new}} - t_1 - t_3 < 0. \quad (46)$$

Thus, $\text{Delay}_{\text{new}} < \text{Delay}_{\text{original}}$, which means choosing the node with the least number of interfering nodes to replace the

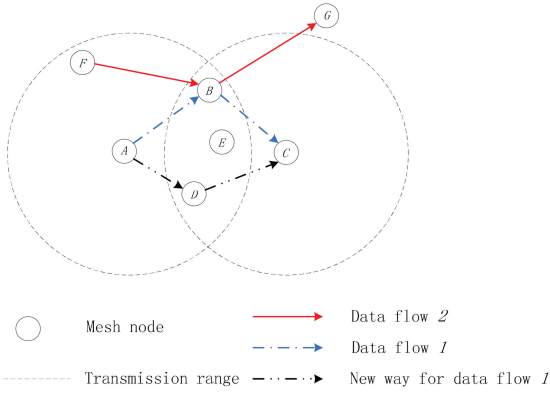


Fig. 4. Example of selecting the replacement node.

original common node can avoid congestion at the common node, balance load, and finally reduce delay.

Furthermore, if other nodes that are not with the least number of interfering nodes are chosen, the delay of the adjacent two links at the replacement node (denoted as $\text{Delay}_{\text{other}}$) is

$$\text{Delay}_{\text{other}} = t_1^{\text{other}} + t_2. \quad (47)$$

Then, the relationship between $\text{Delay}_{\text{new}}$ and $\text{Delay}_{\text{other}}$ is

$$\begin{aligned} \text{Delay}_{\text{new}} - \text{Delay}_{\text{other}} &= \text{Delay}_{\text{min}} - \text{Delay}_{\text{other}} \\ &= (t_1^{\text{min}} + t_2) - (t_1^{\text{other}} + t_2) = t_1^{\text{min}} - t_1^{\text{other}} < 0. \end{aligned} \quad (48)$$

Therefore, choosing the node with the least number of interfering nodes as the replacement node can bring the least delay. A simple example of this method is shown in Fig. 4.

There are two data flows in Fig. 4, and node B is the original common node used by both data flows 1 and 2 . Congestion and long queuing delay will then occur at node B . To avoid this congestion, the improved algorithm of DIAR will choose another node to transmit data flow 1 . Besides node B , nodes E and D are also within the transmission range of nodes A and C , so nodes E and D can be used to replace node B . Assume that the numbers of interfering nodes of nodes B , E , and D (i.e., n_B , n_E , n_D) are 5, 6, and 3, so node D will be selected to replace node B to transmit data flow 1 as node D has the least number of interfering nodes.

D. Cost of DIAR

Considering delay, ETT is a very popular and common method to obtain delay in the routing problem, and it is used as the routing metric in the relevant reference [9]. Therefore, we compare the cost of DIAR and ETT.

1) *Cost in Space*: Unlike ETT, DIAR no longer needs to broadcast extra probe packets when obtaining network conditions. Thus, the cost of extra probe packets and wireless resource is saved. However, to enable DIAR to find the neighbor with the least number of interfering nodes, DIAR needs some storage space in the existing Hello packets to record the number of interfering nodes. Besides, the network manager needs to store the topology showing neighboring and interfering relationships among nodes.

As ETT uses the technique of broadcasting probe packet pairs to evaluate delay [10], the cost in space of ETT in the whole network (denoted as O_{ETT}) is

$$O_{\text{ETT}} = (O_1 + O_2)n_a \cdot m \quad (49)$$

where O_1 is the cost of sending small probe packet, which is 137 bytes. O_2 is the cost of sending a large probe packet, which is 1137 bytes. n_a is the number of active mesh nodes in the network, which need to detect network conditions and select paths. m is the average number of neighboring nodes of active nodes.

DIAR does not have the cost of sending probe packets, but it has more storage cost in Hello packets and network manager. The cost in space to store the number of interfering nodes for one node in each Hello packet and the cost in space of each element in the neighboring and interfering matrix at the network manager are both 4 bytes (denoted as O_e). Therefore, the cost in space of DIAR is

$$O_{\text{DIAR}} = (n + 2n^2)O_e \quad (50)$$

where n is the number of mesh nodes, and n^2 is the matrix size of the neighboring and interfering matrices. From (49) and (50), we can see that the cost in space of DIAR is much less than ETT as long as $\frac{2}{637}(2n^2 + n) < n_a \cdot m$, which is easy to satisfy in a common network.

2) *Computational Complexity*: In DIAR, delay is computed by (19) for each node. The computational complexity depends on obtaining PTF and available bandwidth. According to the derived relationship between PTF and the number of interfering nodes, PTF can be obtained directly based on the number of interfering nodes. Assuming that the number of nodes in a whole network is n , the maximum number of interfering nodes is $n - 2$. When obtaining available bandwidth as (23), the maximum required number of times of adding interference power is $n - 2$. Thus, the computational complexity to calculate the delay for each node is $O(n)$ in general. For ETT, as extra probe packets are used to obtain delay without computing, the computational complexity of ETT is $O(1)$.

VII. SIMULATION EVALUATION

A. Simulation Environment

For the optimization process, our experiments with the 200-iteration GA are taken on a processor Core i7-6700 at 3.4 GHz. The time to reach a good solution is around several seconds. This time can be significantly saved further by using a machine with strong computation ability like a GPU [40]. For the network performance, simulation through NS3 [41] has been implemented. Mesh routers and clients are deployed in an area of $1000 \text{ m} \times 1000 \text{ m}$. Mesh routers are equipped with three radio interfaces, and mesh clients are with a single radio interface. Each radio interface is banded with an orthogonal channel. The packet transmission rate and the number of origin-destination data flows are changed during the simulation. As emergency and rural communications are very important applications of the WMN, instant messaging like VoIP and streaming multimedia communication is used. These applications rely on UDP, so UDP is used as the transport layer protocol. The data flows are constant-bit-rate flows arriving uniformly with random sources and destinations, including both short and long ones. The detailed simulation parameters are shown in Table II.

B. Performance Metrics

Average packet loss rate, average delay, and average network throughput are used as the performance metrics to evaluate the network performance.

TABLE II
SIMULATION PARAMETERS

Simulation Parameters	Values
Simulation time	100s
Traffic type	UDP
Packet size	1024 bytes
Number of mesh routers	25
Number of mesh clients	50
Number of radio interfaces in each router	3
Number of channels in each router	3
Number of radio interfaces in each client	1
Number of channels in each client	1
Transmission range	250m
Interference range	550m
Antenna	Omnidirectional

1) Average packet loss rate (denoted as PI)

$$PI = \frac{N_{\text{loss}}}{N_{\text{total}}} = \frac{N_{\text{total}} - N_{\text{received}}}{N_{\text{total}}} \quad (51)$$

where N_{loss} is the number of packets which are lost during transmission, N_{total} is the total number of packets sent by the source, and N_{received} is the number of packets received successfully by the destination.

2) Average delay (denoted as D)

$$D = \frac{\sum_{i=1}^m D_i}{m} \quad (52)$$

where D_i is the delay cost by transmitting packet i and m is the total number of packets received by the destination in success.

3) Average network throughput (denoted as Th)

$$Th = \frac{N_{\text{received}} \cdot \text{Byte} \cdot 8}{(T_{\text{end}} - T_{\text{start}}) \cdot 1024} \quad (53)$$

where N_{received} is the number of packets received by the destination, Byte is the number of bytes contained in one data packet, T_{end} is the time when the last packet is received successfully, and T_{start} is the time of starting sending the first data packet. The unit of average network throughput is kb/s.

C. Simulation Results and Analyses

To evaluate the performance of DIAR roundly, the rate of sending data packets, the number of end-to-end data flows, and the length of paths are changed. The performance of DIAR, DIAR with basic GA (i.e., DIAR_basic_GA), search-based routing [9], DIAR with ETX (i.e., DIAR_ETX), and routing method with minimized hop count (minHop) is evaluated and compared. DIAR_basic_GA has the same mathematical optimization model with DIAR, but uses the traditional GA rather than the improved one to solve the optimization problem. The minHop is the most widely used routing method, and it is always used as the benchmark [39], [42]. ETX is another popular routing metric to evaluate link condition. To evaluate the performance of the objective in DIAR, we change the objective of DIAR into ETX and still use the improved GA to solve the routing problem. Besides, according to the relevance and timeliness, the search-based routing proposed in [9] is also compared. When the number of data flows is 4, the obtained performance results with different packet transmission rate are shown and compared in Fig. 5(a)–(c). The figure shows that DIAR achieves lower average packet loss rate, lower average delay, and higher network throughput. With the increase of the packet transmission rate, the advantage of DIAR is more significant, because DIAR can balance load and congestion

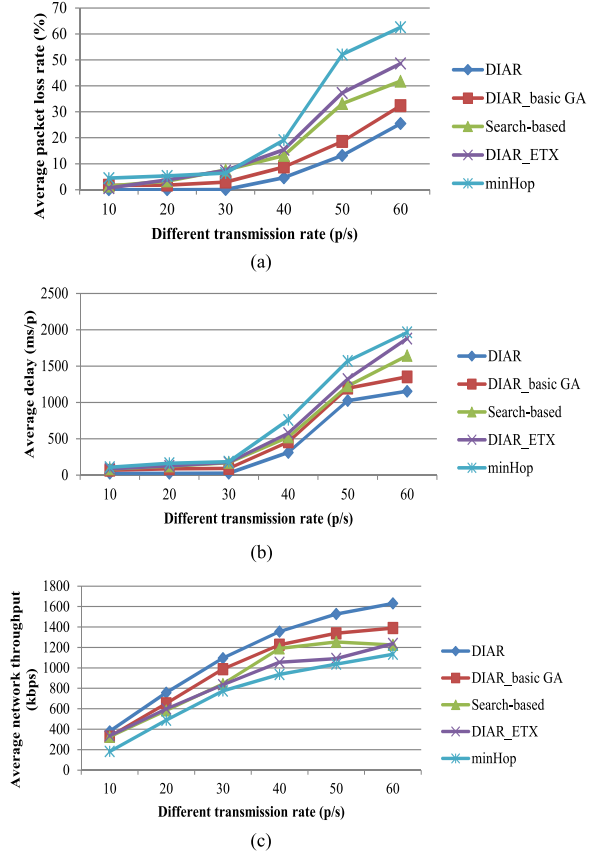


Fig. 5. (a) Average packet loss rate, (b) average delay, and (c) average network throughput with different transmission rates under four data flows.

effectively. DIAR improves the basic GA and avoids the congestion caused by using one common node to serve multiple data flows. Besides, DIAR considers the interference and dynamic network condition during the process of selecting routes, which can help it choose the routes to gain better whole network performance. The search-based routing considers ETT in addition to hop count, so it also takes network condition into account and gets better performance than minHop. However, ETT neglects the backoff delay and brings high overhead. In addition, the search-based routing detects ETT of all links at the beginning of optimization and assumes they are static during selecting paths. When different links are chosen to transmit packets in route discovery process, the ETT values in the network will actually change. They should be evaluated in each iteration of the GA, but the search-based routing does not consider this influence during the process of optimization, which makes it less accurate and its performance worse than DIAR. ETX also considers the packet loss condition, but it neglects bandwidth. Then, the performance of DIAR_ETX is generally a bit worse than ETT. The network performance by using DIAR is 23.1%, 37.8%, 43.7%, and 55.1% better than that of DIAR_basic GA, Search-based, DIAR_ETX, and minHop in average.

Similarly, when the number of data flows is 7, the obtained performance results with different packet transmission rate are shown and compared in Fig. 6(a)–(c). The figure shows that DIAR achieves better network performance in terms of average packet loss rate, average delay, and average network throughput, which is similar to the performance when the number of data flows is four. When the number of data flows increases, the

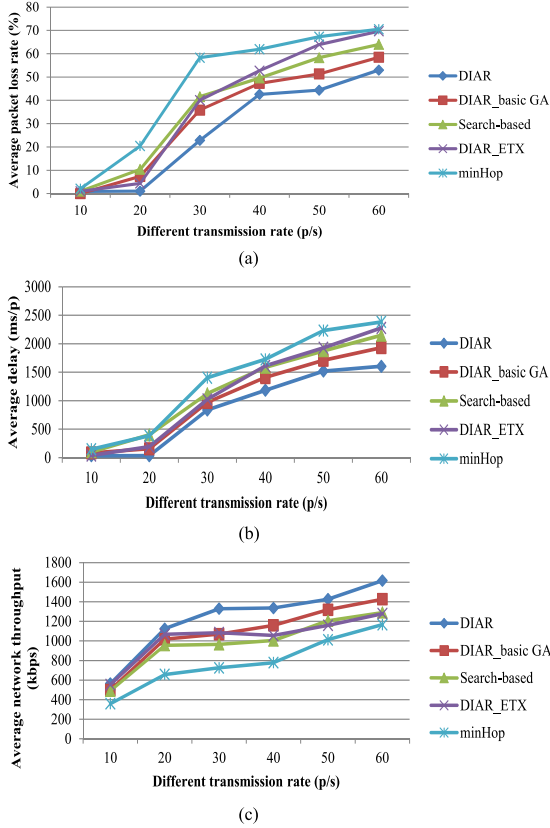


Fig. 6. (a) Average packet loss rate, (b) average delay, and (c) average network throughput with different transmission rates under seven data flows.

probability of congestion will increase. Then, the gap between the DIAR and other three routing methods is wider and the advantage of DIAR is more significant, because it can avoid using the same node too often. Instead, another node with the least number of interfering nodes will be selected as the replacement node. The heavy load can be balanced in this case. In addition, based on the process of optimization, DIAR can select paths with less delay, less interference, less PTF, and larger available bandwidth, which brings better network performance. When changing the objective of DIAR into ETX, the performance will be worse than DIAR because ETX neglects bandwidth, delay, and interference. However, its performance is sometimes better than others because it still uses the improved GA, which can balance load to solve the routing problem. The load-balance solution helps to cover the shortage of ETX in DIAR_ETX to some extent. minHop does not consider any network factors and only selects the path with least hop count to transmit packets, which may cause heavy load and congestion. Thus, minHop gets worst performance. Although the search-based routing considers both hop count and ETT, the relationship of them is product. The factor of hop count can influence the selection of routes in large extent. The search-based routing tends to still choose ways with less hop count when the sums of ETT on diverse routes are not so different. Congested ways with less hop count and long delay may be chosen. The network condition of ETT is not predominant in search-based routing. Therefore, DIAR performs better than search-based routing. The performance of DIAR is 16.2%, 26.6%, 24.9%, and 44.9% better compared to DIAR_basic GA, Search-based, DIAR_ETX, and minHop, respectively.

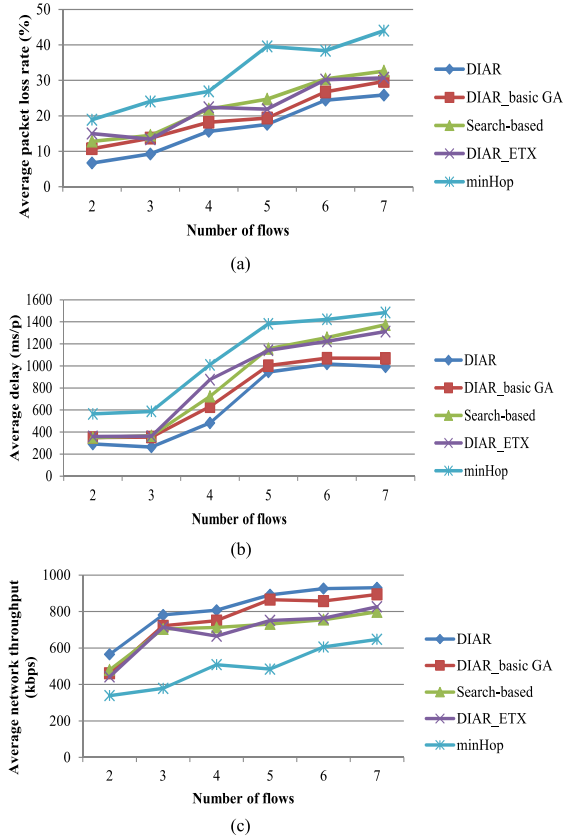


Fig. 7. (a) Average packet loss, (b) average delay, and (c) average network throughput with different number of flows.

When the packet transmission rate is 40 packets per second, the obtained performance results with different number of random end-to-end data flows are shown and compared in Fig. 7(a)–(c). As shown in the figure, with different number of data flows, DIAR can always get better network performance. When the number of data flows increases, the interfering relationships in the network will be more complicated. Large interference will bring bad network performance. minHop does not consider interference, so its performance is worst and fluctuating. ETX considers PTF, but interference and bandwidth are also neglected. However, the load-balance feature of the solution helps DIAR_ETX bring not bad performance with the increasing amount of flows. For DIAR, after establishing the relationship between delay, PTF, and the number of interfering nodes, it can easily obtain delay with low cost and effectively choose the path with better condition. It can also evaluate the dynamic network condition of different path solutions in iterative optimization, which the search-based routing cannot do. DIAR can find paths for several given data flows as a whole and considers both transmission and backoff delay. Thus, DIAR can improve the whole network performance globally and effectively. The network performance of DIAR is 11.5%, 22.7%, 22.6%, and 50.6% better compared to DIAR_basic GA, Search-based, DIAR_ETX, and minHop, respectively, on average.

VIII. CONCLUSION

Effective paths to serve data flows can improve network performance dramatically. Therefore, a good method of selecting paths with high quality is essential. DIAR builds a routing

optimization model to minimize the end-to-end delay. In the process of evaluating delay, the bandwidth, interference, and PTF are considered at the same time. The relationship between delay and the number of interfering nodes is first built in DIAR, so DIAR can obtain delay with low cost. In order to solve the optimization problem more effectively, an improved GA is proposed in DIAR. The improved GA is more adaptive to solve the routing problem. When multiple data flows use a common node, this common node will be replaced by another neighbor node with the least number of interfering nodes. In this way, congestion can be avoided, and load can be balanced effectively. Besides, because the solution of routing can change the interference relationship and network condition, the network condition is, therefore, dynamic. DIAR considers such a dynamic condition and evaluates the network condition every time of finding routes in the iterative optimization. The simulation results through NS3 show that DIAR can achieve better network performance in different cases.

REFERENCES

- [1] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," *Comput. Netw.*, vol. 47, no. 4, pp. 445–487, Mar. 2005.
- [2] S. Paris, C. Nita-Rotaru, F. Martignon, and A. Capone, "Cross-layer metrics for reliable routing in wireless mesh networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 3, pp. 1003–1016, Jun. 2013.
- [3] D. Benyamina, A. Hafid, and M. Gendreau, "Wireless mesh networks design—A survey," *IEEE Commun. Surv. Tut.*, vol. 14, no. 2, pp. 299–310, Second Quarter, 2012.
- [4] S. K. Dhurandher, I. Woungang, M. S. Obaidat, K. Kumar, M. Joshi, and M. Verma, "A distributed adaptive admission control scheme for multimedia wireless mesh networks," *IEEE Syst. J.*, vol. 9, no. 2, pp. 595–604, Feb. 2014.
- [5] S. Khan, K. K. Loo, N. Mast, and T. Naeem, "SRPM: Secure routing protocol for IEEE 802.11 infrastructure based wireless mesh networks," *J. Netw. Syst. Manage.*, vol. 18, no. 2, pp. 190–209, Jun. 2010.
- [6] A. Al-Saadi, R. Setchi, Y. Hicks, and S. M. Allen, "Routing protocol for heterogeneous wireless mesh networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9773–9786, Dec. 2016.
- [7] M. Boushaba, A. Hafid, and M. Gendreau, "Source-based routing in wireless mesh networks," *IEEE Syst. J.*, vol. 10, no. 1, pp. 262–270, May 2014.
- [8] M. Srinivas and L. M. Patnaik, "Genetic algorithms: A survey," *Comput.*, vol. 27, no. 6, pp. 17–26, Jun. 1994.
- [9] K. Mahmood, B. Nazir, I. A. Khan, and N. Shah, "Search-based routing in wireless mesh network," *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, Dec. 2017, Art. no. 36.
- [10] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proc. 10th ACM Int. Conf. Mobile Comput. Netw.*, Sep. 2004, pp. 114–128.
- [11] S. Yussof and O. H. See, "A robust GA-based QoS routing algorithm for solving multi-constrained path problem," *J. Comput.*, vol. 5, no. 9, pp. 1322–1334, Sep. 2010.
- [12] C. H. Liu, A. Gkelias, and K. K. Leung, "A cross-layer framework of QoS routing and distributed scheduling for mesh networks," in *Proc. IEEE Veh. Technol. Conf.*, May 2008, pp. 2193–2197.
- [13] R. Murugeswari, S. Radhakrishnan, and D. Devaraj, "A multi-objective evolutionary algorithm based QoS routing in wireless mesh networks," *Appl. Soft Comput.*, vol. 40, pp. 517–525, Mar. 2016.
- [14] D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Wireless Netw.*, vol. 11, no. 4, pp. 419–434, Jul. 2005.
- [15] E. Alotaibi, V. Ramamurthi, M. Batayneh, and B. Mukherjee, "Interference-aware routing for multi-hop wireless mesh networks," *Comput. Commun.*, vol. 33, no. 16, pp. 1961–1971, Oct. 2010.
- [16] T. Meng, F. Wu, Z. Yang, G. Chen, and A. V. Vasilakos, "Spatial reusability-aware routing in multi-hop wireless networks," *IEEE Trans. Comput.*, vol. 65, no. 1, pp. 244–255, Jan. 2016.
- [17] W. Shi, E. Wang, Z. Li, Z. Wang, and S. Wang, "A load-balance and interference-aware routing algorithm for multicast in wireless mesh network," in *Proc. Int. Conf. Inf. Commun. Technol. Convergence*, Oct. 2016, pp. 206–210.
- [18] S. Middha, and P. Singh, "Channel assignment with greedy algorithm for wireless mesh network," *Int. J. Modern Comput. Sci. Appl.*, vol. 3, no. 3, pp. 21–26, Aug. 2015.
- [19] Y. Wang, G. Cong, G. Song, and K. Xie, "Community-based greedy algorithm for mining top-k influential nodes in mobile social networks," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2010, pp. 1039–1048.
- [20] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 protocol: Design and performance evaluation of an adaptive backoff mechanism," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 9, pp. 1774–1786, Sep. 2000.
- [21] V. A. Siris, G. Stamatakis, and E. Tragos, "A simple end-to-end throughput model for 802.11 multi-radio multi-rate wireless mesh networks," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 635–637, Apr. 2011.
- [22] E. Ancillotti, R. Bruno, M. Conti, and A. Pinizzotto, "Load-aware routing in mesh networks: Models, algorithms and experimentation," *Comput. Commun.*, vol. 34, no. 8, pp. 948–961, Jun. 2011.
- [23] X. Cao, L. Liu, W. Shen, and Y. Cheng, "Distributed scheduling and delay-aware routing in multihop MR-MC wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6330–6342, Aug. 2016.
- [24] H. Jiang, L. Zheng, Y. Liu, and M. Zhang, "Multi-constrained QoS routing optimization of wireless mesh network based on hybrid genetic algorithm," in *Proc. Int. Conf. Intell. Comput. Integr. Syst.*, Oct. 2010, pp. 862–865.
- [25] M. Dorigo, and M. Birattari, "Ant colony optimization," in *Encyclopedia of Machine Learning*. Boston, MA, USA: Springer, 2011, pp. 36–39.
- [26] V. Sarasvathi, N. C. S. Iyengar, and S. Saha, "QoS guaranteed intelligent routing using hybrid PSO-GA in wireless mesh networks," *Cybern. Inf. Technol.*, vol. 15, no. 1, pp. 69–83, Mar. 2015.
- [27] Y. Shi, "Particle swarm optimization: Developments, applications and resources," in *Proc. Congr. Evol. Comput.*, May 2001, vol. 1, pp. 81–86.
- [28] V. V. Mandhare, V. R. Thool, and R. R. Manthalkar, "QoS Routing enhancement using metaheuristic approach in mobile ad-hoc network," *Comput. Netw.*, vol. 110, pp. 180–191, Dec. 2016.
- [29] H. Huang, P. Li, S. Guo, and W. Zhuang, "Software-defined wireless mesh networks: Architecture and traffic orchestration," *IEEE Netw.*, vol. 29, no. 4, pp. 24–30, Jul. 2015.
- [30] O. Dousse, M. Franceschetti, N. Macris, R. Meester, and P. Thiran, "Percolation in the signal to interference ratio graph," *J. Appl. Probab.*, vol. 43, no. 2, pp. 552–562, Jun. 2006.
- [31] A. U. Chaudhry, R. H. Hafez, O. Aboul-Magd, and S. A. Mahmoud, "Throughput improvement in multi-radio multi-channel 802.11a-based wireless mesh networks," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2010, pp. 1–5.
- [32] R. Morsi, D. S. Michalopoulos, and R. Schober, "Multiuser scheduling schemes for simultaneous wireless information and power transfer over fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 1967–1982, Apr. 2015.
- [33] M. A. Mohammed, M. K. A. Hamed, R. I. Ghani, S. A. Mostafa, M. S. Ahmad, and D. A. Ibrahim, "Solving vehicle routing problem by using improved genetic algorithm for optimal solution," *J. Comput. Sci.*, vol. 21, pp. 255–262, Jul. 2017.
- [34] H. Li, Y. Cheng, C. Zhou, and W. Zhuang, "Minimizing end-to-end delay: A novel routing metric for multi-radio wireless mesh networks," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 46–54.
- [35] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, Aug. 2002.
- [36] N. Chakchouk, and B. Hamdaoui, "Traffic and interference aware scheduling for multiradio multichannel wireless mesh networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 555–565, Feb. 2011.
- [37] J. Kim, K. Kim, and J. Lee, "Energy-efficient relay selection of cooperative HARQ based on the number of transmissions over Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 610–621, Jan. 2017.
- [38] Y. Hu, A. Schmeink, and J. Gross, "Blocklength-limited performance of relaying under quasi-static Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4548–4558, Jul. 2016.
- [39] C. Lozano-Garzon, M. Camelo, P. Vilà, and Y. Donoso, "A multi-objective routing algorithm for wireless mesh network in a smart cities environment," *J. Netw.*, vol. 10, no. 1, pp. 60–69, Feb. 2015.
- [40] D. Akgün and P. Erdomus, "GPU accelerated training of image convolution filter weights using genetic algorithms," *Appl. Soft Comput.*, vol. 30, pp. 585–594, May 2015.
- [41] NS3. Nov. 9, 2018. [Online]. Available: <http://www.nsnam.org/>
- [42] M. Camelo, C. Omana, and H. Castro, "QoS routing algorithms based on multi-objective optimization for mesh networks," *IEEE Latin Amer. Trans.*, vol. 9, no. 5, pp. 875–881, Sep. 2011.