# A Low-Latency Interference Coordinated Routing for Wireless Multi-hop Networks

Jianming Cheng, Pei Yang, Keivan Navaie, *Senior Member, IEEE*,
Qiang Ni, *Senior Member, IEEE*, and Hongwen Yang, *Member, IEEE*

Abstract— **Recently, there has been an increasing interest in exploiting interference cancelation to support multiple adjacent concurrent transmissions instead of avoiding interference through scheduling. In line with these efforts, this paper propose an interference coordinated routing (ICR) scheme for wireless multi-hop networks to achieve more transmission concurrence, and thus lower the end-to-end delay. The proposed ICR scheme firstly constructs an initial path by the interference-aware routing algorithm, which captures the end-to-end latency and spatial resource cost as the routing metrics. Then, to analyze the feasibility of concurrent transmission for a given link set, we consider the interference coordination and formulate the concurrent transmission of multiple links as a linear programming (LP) problem. The solution to the LP problem indicates the power allocation. Finally, a distributed guard zone based selection (GBS) algorithm is further proposed to iteratively explore the maximum feasible link set for each time slot. The selected links are simultaneously active for packet transmission with the allocated power in the current time slot, and the remaining links will be put off to the next. Simulation results confirm that ICR reduces the end-to-end delay by 9.16% to 73.82%, and promotes better transmission concurrence compared with the existing schemes.**

Index Terms— **Wireless multi-hop networks, routing, successive interference cancellation (SIC), superposition coding (SC), end-to-end delay.**
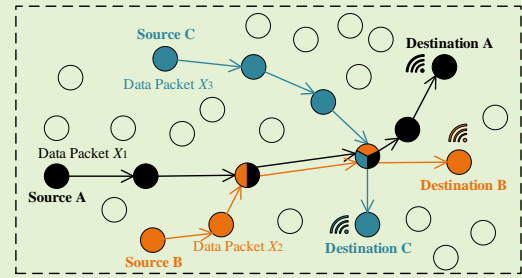
I. INTRODUCTION

IN recent years, the rapid development of the Internet of Things (IoT) and 5G network has enabled the wireless self-organizing multi-hop network to support wider applications in the harshest of environments, such as volcanoes, hurricane-affected regions, and underground mines [1]–[5]. In these challenging environments, the parallel data packet transmissions from multiple nodes become a common scenario, and mutual interference is the leading cause of congestion, which further results in high end-to-end latency. Therefore, investigating interference-aware low-latency communication is a continuing concern within these wireless scenarios. Most IoT enabled devices are equipped with a wireless transceiver to exchange data with other neighboring nodes, and, when necessary, to relay packets via neighboring nodes to destinations that are not within direct communications of each other. Accordingly, the main challenge faced by many researchers is how to design an interference-aware low-latency routing scheme for wireless multi-hop scenarios [6]–[8].

Interference is a fundamental impediment to the end-to-end delay in a wireless network since the wireless channel is a shared and thus interference-limited medium. In a local area, e.g., the center area of network, the accumulated interference among multiple links may be too severe to transmit data simultaneously which gives rise to the hot spot problem [9]. Existing routing protocols [10]–[16] generally added interference awareness into routing protocols to avoid routes passing through the hot spot of interference, increase the transmission concurrence and improve the network throughput. However, these routing protocols come at the cost of increased end-to-end delay. Some recent studies [17], [18] have suggested that the cross-layer design could effectively mitigate the interference by combining the routing process with the scheduling strategies in the medium access control (MAC) layer and interference cancelation techniques in the physical layer. Time-division-multiple-access (TDMA) is a MAC mechanism commonly used in wireless sensor networks or wireless ad hoc networks, where the MAC time is divided into time slots, and links are active for data transmission only in their assigned time slots. Considering that end-to-end multi-hop transmission costs multiple time slots, most of the previous cross-layer low-latency routing protocols [19]–[23] focused on designing the optimal scheduling for finding an optimal simultaneous link set in each time slot without causing strong interference in the network, so that the activation of interfering links can be well-managed to achieve optimal spatial reuse. However, these

studies failed to resolve the hot spot problem. Consequently, link scheduling may cause a long end-to-end delay.

Instead of avoiding the interference through scheduling, some recent studies have shown an increasing interest in exploiting the interference coordination to allow multiple adjacent concurrent transmissions to coexist [24], [25]. Successive interference cancelation (SIC) and superposition coding (SC) are two well-known physical layer interference coordination techniques. The former enables simultaneous unicast transmissions from multiple senders to a single receiver, and the latter enables simultaneous unicast transmissions from a single sender to multiple receivers.

A considerable amount of literature has been published on interference coordination [26]. Among these studies, only a couple of papers [27]–[36] focused on applying SIC or SC to the scheduling and routing scheme so far. A source routing with SIC and SC applied (S3) protocol was proposed in [27]. It introduced the interference aware expected transmission time as the routing metric, but the effect that interference coordination enables the receiver to partially cancel the interfering signals was not analyzed. Later, some network graph models (the conflict set graph [28] and simultaneity graph [29], [30]) were applied to characterize the effect of SIC and evaluate the link interference for scheduling. In [31]–[35], SIC was further formulated in the cross-layer (physical, link and network) optimization frameworks, and it has been shown that substantial performance gains can be obtained by exploiting interference. In this sense, the benefits of SIC was analyzed and the SIC-aware routing (SAR) scheme was proposed in [36]. SAR constructed the end-to-end path based on the spatial resource cost and prevented the routing path from passing through the hot spot area. Meanwhile, it defined the SIC-able condition to discover the communication links with potential SIC opportunities, thereby improving the end-to-end throughput and mitigating the hot spot problem. However, the existing routing designs only considered the interference cancelation case that each receiver can cancel at most one signal, which was not efficient for the concurrence of multiple data packets. When multiple parallel data packet transmissions happen, SIC should be applied and the order of cancelation needs to be analyzed.

Clearly, efficient link scheduling, together with interference coordination, contributes to promoting better spatial reuse as well as transmission concurrence, which can be used in a wireless multi-hop network to further lower end-to-end delay. Motivated by these, this paper considers a wireless network with multiple end-to-end data packets and proposes an interference coordinated routing (ICR) scheme. In contrast to the existing schemes, the *importance and originality* of this study can be summarized as follows:

Firstly, an interference-aware routing is designed to construct the end-to-end initial path, with the end-to-end latency and spatial resource cost captured as the routing metrics. Then, this paper analyzes the feasibility of concurrent transmissions for a given link set. The concurrent transmissions of multiple links satisfying the half-duplex condition are formulated as a linear programming (LP) problem, and the power allocation obtained from solving LP problem is checked by the mini-

mum signal-to-interference-plus-noise ratio (SINR) condition. Finally, this paper designs a distributed guard zone based selection (GBS) algorithm to iteratively explore the maximum feasible link set for each time slot. The selected links will be active for transmission according to the allocated power in the current time slot, and the data packets of the remaining links will be put off to the next time slot. By running the proposed GBS algorithm slot by slot, the interference-aware scheduling for multiple paths can be achieved in a compact manner.

In a nutshell, the proposed ICR scheme achieves a distributed cross-layer design (routing in network layer, link scheduling in MAC layer, and interference-aware power allocation in physical layer). Extensive simulations are conducted for performance evaluation, which demonstrates the effectiveness of the proposed ICR scheme in terms of improving the transmission concurrence and thus reducing the end-to-end delay.

The remainder of this paper is organized as the following. Section II illustrates the system model and the interference coordination methods: SIC and SC. Section III presents the proposed ICR scheme. Simulation results are provided in Section IV. Finally, conclusions are drawn in Section V.

II. SYSTEM MODEL AND INTERFERENCE COORDINATION

A. Network Model

We consider a network consisting of N stationary nodes arbitrarily distributed in an area, where each node is equipped with a single omnidirectional antenna and operates in half-duplex mode. These nodes can self-organize to form a multi-hop network. The whole network can be defined as a connected undirected graph: namely, $G = (\mathcal{V}, \mathcal{P})$, where $\mathcal{V}$ is the set of nodes and $|\mathcal{V}| = N$ is the number of nodes. $\mathcal{P}$ denotes the set of all bi-directional wireless communication links between pairs of nodes. Each node i can adjust its transmission power P_i up to the maximum limit $P_{\max}$. In case a node $i \in \mathcal{V}$ transmits data X with the maximum power $P_{\max}$, and another node $j \in \mathcal{V}$ can successfully decode X without the aid of any other nodes, we say $\mathcal{P}_{i,j} \in \mathcal{P}$ exists.

Consider the Poisson traffic model as the packet generation model of the network, where the packet generation process for each node per time unit follows the Poisson distribution. Thus the total number of packets in the network is changing with time, and every time a packet is newly generated, the source node will initiate a fast routing process to construct an end-to-end path towards the related destination node. Assuming that all the packets generated in the network have the same bit length, β is the required SNR threshold for a successful packet transmission. According to the threshold-based reception model, the *Neighbor Set* of node i, $\mathcal{N}(i)$, consists of all reachable nodes of i:

$$\mathcal{N}(i) = \left\{ j \in \mathcal{V} \,\middle|\, \frac{P_{\max} g_{i,j}}{\sigma^2} \geq \beta \right\}, \tag{1}$$

where σ^2 is the power level of the ambient noise, and $g_{i,j}$ is the channel gain between node i and j, which contains the effect of both large-scale path loss and small-scale fading. We assume that all links in $\mathcal{P}$ are subjected to independent

Rayleigh fading with average strength $\mathbb{E}[g_{i,j}] = d_{i,j}^{-\alpha}$, where $d_{i,j}$ denotes the Euclidean distance between any node i and j, and α is the path-loss exponent. We assume that the network topology $G = (\mathcal{V}, \mathcal{P})$ and channel gain $\{g_{i,j}\}$ keep unchanged during the end-to-end transmission of each packet. In practice, the network topology, especially the channel gain may change slowly. In this case, the reactive routing ways [39] can be used to adapt to such change.

Similar to the existing research, see, e.g., [31], we assume that the data scheduling or transmission occurs in time-slot basis so that the slot length is the time required to transmit one packet over a link. Meanwhile, each node stores the packets generated or successfully decoded in its buffer which is assumed large enough to contain these packets. During each time slot, only the selected packets can be transmitted over the related links, while other packets wait for the next available time slot, i.e., these packets experience the queuing delay. Let us assume that all the nodes are synchronized. The end-to-end delay for a data packet is calculated from the moment this packet is generated at the source to the moment it has been decoded successfully at the destination, including transmission delay over the multi-hop links and the possible queuing delay at source and each intermediate relay node. Moreover, we assume that, the control messages are transmitted reliably with the negligible cost due to the fact that the length in bits of these messages is very small compared to the data packet [10].

Similar to [37], it suggests that nodes always have a means to acquire the channel state information (CSI) of their neighbors. Consider two links $\mathcal{P}_{i,j}$ and $\mathcal{P}_{m,j}$, which are active simultaneously. Under the above-mentioned reception model, the reception of link $\mathcal{P}_{i,j}$ is successful if the received SINR

$$\frac{P_{\max} \cdot g_{i,j}}{P_{\max} \cdot g_{m,j} + \sigma^2} \geq \beta. \tag{2}$$

To ensure (2), the following inequality must be guaranteed:

$$g_{m,j} < \beta^{-1} g_{i,j}, \tag{3}$$

or node j should first cancel the interference from node m. Therefore, we define the *guard zone* of link $\mathcal{P}_{i,j}$, denoted by $A_{i,j}$, as the set of nodes which violates (3). Specifically,

$$A_{i,j} = \left\{ m \in \mathcal{V} \, \middle| \, g_{m,j} \geq \beta^{-1} g_{i,j} \right\}, \tag{4}$$

and $S_{i,j} = |A_{i,j}|$ represents the number of nodes inside $A_{i,j}$. For example, as shown in Fig. 1a, $A_{i,j} = \{i, j, m\}$ and $A_{m,j} = \{m, j\}$, so $S_{i,j} = 3$ and $S_{m,j} = 2$. It can be observed from (4) that for $\beta = 1$, $A_{i,j}$ denotes a set of nodes which have larger channel gains than node i. For $\beta \to 0$, (2) always holds regardless of the interference from any node $m \in \mathcal{V} \backslash i$, in this case, $A_{i,j} = \{i, j\}$. Oppositely, for $\beta \to \infty$, $A_{i,j} = \mathcal{V}$.

While $\mathcal{P}_{i,j}$ is active, any node in $A_{i,j} \backslash i$ will either be prohibited from transmitting or it must satisfy a set of constraints which we will state later. On the other hand, (2) indicates that a single interferer from $A_{i,j}^{\mathrm{C}} = \mathcal{V} \backslash A_{i,j}$ is acceptable. Although multiple interferences from $A_{i,j}^{\mathrm{C}}$ may degrade the SINR of link $\mathcal{P}_{i,j}$ to a level below β, this can be avoided by appropriate power allocation and link scheduling. In general, for the path-finding process of a newly-generated packet, while $\mathcal{P}_{i,j}$ is active, we have more freedom to select intermediate nodes

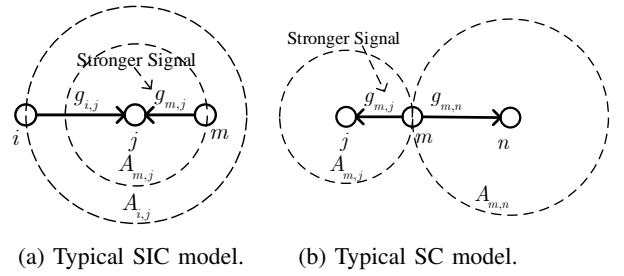(a) Typical SIC model.   (b) Typical SC model.

Fig. 1: Illustrations of the guard zone and interference coordination. In order to show the concept in an intuitive manner, we assume $\beta = 1$ and ignore the small-scale fading in this figure, thus the order of signal strength of links is the reverse order of Euclidean link distances.

from $A_{i,j}^{\mathrm{C}}$ than from $A_{i,j}$. Thus, to some extent, the size of $A_{i,j}$ can reflect the spatial resource occupied by link $\mathcal{P}_{i,j}$.

For the interference from any transmitter $m \in A_{i,j} \backslash i$, interference coordination can be used to enable the concurrent transmissions by canceling this interference. We define the *degree of guard zone* $D_{i,j}$ for link $\mathcal{P}_{i,j}$ as the number of upcoming transmitters inside $A_{i,j}$, and $1 \leq D_{i,j} \leq S_{i,j}$. Take three upcoming links $\{\mathcal{P}_{i,j}, \mathcal{P}_{m,j}, \mathcal{P}_{m,n}\}$ in Fig. 1 as an example, it can be observed that $D_{i,j} = 2$, $D_{m,j} = 1$ and $D_{m,n} = 1$. For $\mathcal{P}_{i,j}$, $D_{i,j}$ indicates the number of interference to be canceled before the desired signal at node j, and implies the level of temporary strong interference on $\mathcal{P}_{i,j}$.

B. Interference Coordination: SIC and SC

If there are multiple transmitters in the area of $A_{i,j}$, under the traditional reception model, only one intended transmitter i will be allowed to transmit. Concurrent transmissions to the same receiver will lead to a collision and the reception of link $\mathcal{P}_{i,j}$ will be prevented by the strong interfering signal from other transmitters. In contrast, a SIC receiver is capable of receiving from multiple transmitters at the same time and thus can substantially increase throughput in the network. Fig. 1a illustrates the typical SIC scenario where node j is receiving from two concurrent transmitters i and m, $g_{i,j} < g_{m,j}$. The stronger signal from node m, no matter whether it is the desired signal or not, can be firstly successfully decoded only if its SINR is no less than β, i.e., $\frac{P_m \cdot g_{m,j}}{P_i \cdot g_{i,j} + \sigma^2} \geq \beta$. Then this signal will be subtracted from the aggregate signal, and the desired signal from node i can be decoded successfully if and only if $\frac{P_i \cdot g_{i,j}}{\sigma^2} \geq \beta$.

Correspondingly, if M links share a common transmitter, under the traditional transmission model, these links will be scheduled for transmission in the different time slot. A SC transmitter can split its transmission power into M parts, for simultaneous transmission to M different receivers. The receivers then decode the signal destined for them by SIC. As shown in Fig. 1b, node m simultaneously transmits two signals to j and n, and $g_{m,n} < g_{m,j}$. P_m^j and P_m^n are the allocated power for link $\mathcal{P}_{m,j}$ and $\mathcal{P}_{m,n}$, respectively ($P_m^j + P_m^n = P_m \leq$

$P_{\max}$). When the following SINR conditions are satisfied,

$$
\begin{cases}
\dfrac{P_m^n \cdot g_{m,n}}{P_m^j \cdot g_{m,n} + \sigma^2} \geq \beta \\[2mm]
\dfrac{P_m^n \cdot g_{m,j}}{P_m^j \cdot g_{m,j} + \sigma^2} \geq \beta \\[2mm]
\dfrac{P_m^j \cdot g_{m,j}}{\sigma^2} \geq \beta \\[2mm]
P_m^j + P_m^n \leq P_{\max},
\end{cases}
\tag{5}
$$

the closer receiver j can decode both signals while the farther receiver n can only decode its desired signal.

III. PROPOSED ICR SCHEME

In this section, the proposed ICR scheme is introduced to reduce the end-to-end delay. Suppose that $F-1$ source-destination pairs already exist in the network, and a source-destination pair $\{V_s^*, V_d^*\}$ newly generates[1]. The proposed ICR scheme is unfolded by the following three parts:

1) First, an interference-aware routing algorithm is proposed to construct an initial routing path r^* for $\{V_s^*, V_d^*\}$.
2) Then, for a given link set, the feasible conditions of concurrent transmission are analyzed and formulated as a LP problem. The feasibility can be checked by solving the LP problem.
3) Finally, to explore the maximum feasible link set for each time slot, the interference-aware scheduling for multiple paths is discussed and a guard zone based selection (GBS) algorithm is further proposed.

A. Initial routing

In this subsection, we present an algorithm, named *interference-aware routing algorithm*, to construct an end-to-end initial path. Let us suppose that $r^* = \{r_1, r_2, ..., r_K\}$ is the constructed initial path from V_s^* to V_d^*, $r_1 = V_s^*$, and $r_K = V_d^*$. Here, the end-to-end latency over an initial path depends on the links in this path. The number of links indicates the required latency cost in packet delivery over these links. For example, the packet transmission times over r^* is $K-1$. Therefore, selecting the initial path with a smaller wireless hop count helps in achieving a lower end-to-end latency.

On the other hand, the links from the existing $F-1$ paths can affect the current availability of the links on r^*. Take any link $\mathcal{P}_{i,j} \in r^*$ as an example. In the t-th time slot, if one of links $\{\mathcal{P}_{k,i}, \mathcal{P}_{j,k}\}, k \in \mathcal{V}$ is active, $\mathcal{P}_{i,j}$ will be unavailable because both nodes i and j operate in half-duplex mode. Moreover, accounting for the fact that the SINR threshold β only works for the transmission of one packet, $\mathcal{P}_{i,j}$ cannot simultaneously serve for two paths at the t-th slot, so $\mathcal{P}_{i,j}$ for r^* will be put off if $\mathcal{P}_{i,j}$ for another path becomes active. Overall, apart from the wireless hop count, the interference

[1]For a given source-destination pair $\{V_s, V_d\}$, any newly generated packet is regarded as a new flow with source-destination pair $\{V_s^*, V_d^*\}$ coinciding with $\{V_s, V_d\}$. This new flow may share the same route with previous packet if the overhead of route finding is of primary concern. However, to optimize the performance, each packet should be assigned a new route path which may not be the same one as the previous packet of same source-destination pair due to the dynamic change of interferences.

from the existing paths should also be considered in designing the routing scheme.

Motivated by these, the link cost of $\mathcal{P}_{i,j} \in \mathcal{P}$ is designed as

$$
\mathcal{C}_{i,j} = T_{i,j}^{\mathrm{T}} + T_{i,j}^{\mathrm{I}},
\tag{6}
$$

where $T_{i,j}^T$ represents the latency cost in the packet transmission over the link, and since each data transmission over a link takes exactly one time slot, $T_{i,j}^T = 1$. $T_{i,j}^I$ is the expectation of queuing delay due to the interference from the existing paths.

Suppose that, at some moment t_0, there are $F-1$ existing source-destination paths, where the f-th path has $\eta^f, f = 1, \cdots, F-1$ waiting transmissions, i.e., the f-th path still needs η^f transmissions to finish. At this moment t_0, a new source-destination pair generates and an end-to-end path for this F-th packet needs to be established. The routing algorithm is aware of all existing $F-1$ paths but it cannot foresee the time slots scheduling for the future transmissions of existing paths, since the slot scheduling and power allocation is performed at link and physical layer on a slot-by-slot basis. From the perspective of routing algorithm, any candidate link $\mathcal{P}_{i,j}$ of path F may meet any one of η^f waiting transmissions of path $f = 1, \cdots, F-1$ in any future slot. The probability for meeting a specific waiting transmission of path f is $\frac{1}{\eta^f}$. If this specific waiting transmission meets one of the following three cases: occupying either transmitter i as the receiver or receiver j as the transmitter, or sharing the candidate link $\mathcal{P}_{i,j}$, the transmission of path F passing through $\mathcal{P}_{i,j}$ will highly likely be postponed to the next time slot, leading to an increase of delay by 1 time slot. Considering that there is no loop in any existing end-to-end path, the number of transmissions satisfying any of the above three cases for the f-th path is no more than 1, and thus $T_{i,j}^I$ can be calculated by

$$
T_{i,j}^I = \sum_{f=1}^{F-1} \frac{1}{\eta^f} \cdot \left(x_i^f + y_j^f + z_{i,j}^f \right),
\tag{7}
$$

where $x_i^f, y_j^f, z_{i,j}^f \in \{0, 1\}$ indicates whether transmitter i is occupied as a receiver, receiver j is occupied as a transmitter or link $\mathcal{P}_{i,j}$ is shared by the f-th path, respectively.

It is noted that when $\{V_s^*, V_d^*\}$ is the first source-destination pair, $T_{i,j}^I = 0$ for any link $\mathcal{P}_{i,j} \in \mathcal{P}$. Since there may exist multiple paths with the minimum link cost, to ensure the uniqueness of the initial path, the spatial resource cost S_{r^*} for the path r^* is defined as the second routing metric. For $\mathcal{P}_{i,j} \in \mathcal{P}$, $S_{i,j} = |A_{i,j}|$ is the spatial resource occupied by $A_{i,j}$, and thus S_{r^*} is obtained by

$$
S_{r^*} = \left| \bigcup_{\mathcal{P}_{i,j} \in r^*} A_{i,j} \right|.
\tag{8}
$$

The second routing metric S_{r^*} will be only checked on condition that the end-to-end latency cost cannot guarantee the uniqueness of initial path.

V_s^* initiates the routing discovery and finds the end-to-end initial path with the minimum end-to-end latency and spatial resource cost, i.e.,

$$
r_{V_s^*, V_d^*}^* = \underset{r \in \Omega\left(V_s^*, V_d^*\right)}{\arg\min} \left\{ \sum_{k=1}^{K-1} \mathcal{C}_{r_k, r_{k+1}}, S_r \right\}^2,
\tag{9}
$$

Algorithm 1: Interference-aware routing algorithm

Input: $V_s^*, V_d^*, G = \{\mathcal{V}, \mathcal{P}\}, r^{(f)}, f = 1, \cdots, F-1$.
Output: Initial Path $r_{V_s^*, V_d^*}^*$ for $\{V_s^*, V_d^*\}$.

1 Each node $i \in \mathcal{V}$ calculates $\mathbf{C}_i$ and $\mathbf{S}_i$. $\mathcal{Q} \leftarrow \emptyset$.
2 V_s^* broadcasts RREQ message including $\mathbf{C}_{V_s^*}, \mathbf{S}_{V_s^*}$,
  $\quad \mathcal{Q} \leftarrow \mathcal{Q} \cup V_s^*$.
3 Find $i \leftarrow \underset{i \in \mathcal{V} \setminus \mathcal{Q}}{\arg\min}(\mathcal{C}_{s,i}, S_{s,i})$. $r_{V_s^*, i} = \{V_s^* \rightarrow i\}$.
4 **while** $i \neq V_d^*$ **do**
5 $\quad$ Node i broadcasts RREQ message with $\mathbf{C}_i, \mathbf{S}_i$,
  $\qquad \mathcal{Q} \leftarrow V_s^* \cup i$.
6 $\quad$ For each node $j \in \mathcal{N}(i)$, updates $\{\mathcal{C}_{s,j}, S_{s,j}\}$ by
7 $\quad$ **if** $\mathcal{C}_{s,i} + \mathcal{C}_{i,j} < \mathcal{C}_{s,j}$ **then**
8 $\quad\quad$ $\mathcal{C}_{s,j} \leftarrow (\mathcal{C}_{s,i} + \mathcal{C}_{i,j}), r_{V_s^*, j} = (r_{V_s^*, i} \rightarrow j)$
  $\qquad\quad S_{s,j} \leftarrow S_{r_{V_s^*, j}}$.
9 $\quad$ **end**
10 $\quad$ Find $i \leftarrow \underset{i \in \mathcal{V} \setminus \mathcal{Q}}{\arg\min}(\mathcal{C}_{s,i}, S_{s,i})$.
11 **end**
12 V_d^* sends a unicast RREP message along with $r_{V_s^*, V_d^*}^*$.

---

where $\Omega(V_s^*, V_d^*)$ is the set of all possible routes from V_s^* to V_d^*.

Nodes in the network broadcast the control message to make a fast neighbor discovery to collect the neighborhood information [37]. Through measuring the received signal strength, each node $i \in \mathcal{V}$ maintains two N-dimensional adjacency vectors, $\mathbf{C}_i = \{\mathcal{C}_{i,j}, j \in \mathcal{V}\}$ and $\mathbf{S}_i = \{S_{r_{i,j}}, j \in \mathcal{V}\}$, which records the end-to-end latency and spatial cost vector by (6) and (8), respectively. Denote $\mathcal{Q}$ as the set of nodes which have broadcast the routing request (RREQ) message, the proposed interference-aware routing algorithm can be shown in Algorithm 1.

It is noted that Algorithm 1 is designed based on the typical path-finding process of Ad-hoc On-demand Distance Vector (AODV) routing [39], which contains two steps: the current node i, with the minimum latency and resource cost to the source, broadcasts the RREQ message including its adjacent vectors, $\{\mathbf{C}_i, \mathbf{S}_i\}$, and put i into $\mathcal{Q} = \mathcal{Q} \cup i$; Then any node j received the RREQ message updates its adjacent vectors, $\{\mathbf{C}_j, \mathbf{S}_j\}$, and records the corresponding route, $r_{V_s^*, j}$, accessing V_s^*. We then iterate these two steps until the destination node, V_d^*, becomes the current node, and V_d^* sends a unicast route reply (RREP) message along the reverse of $r_{V_s^*, V_d^*}^*$ to determine the end-to-end path. The proposed algorithm can find an end-to-end initial path between V_s^* and V_d^* based on the minimum end-to-end latency and spatial resource cost. Clearly, this design can reduce the interference to other end-to-end traffic and contribute to the low end-to-end delay.

---

[2]In general, $\arg\min_{i}\{A_i\}$ represents a set of indices corresponding to $\min\{A_i\}$. In this paper, we use $i^* = \arg\min_{i}\{A_i, B_i\}$ to denote the unique index such that $A_{i^*} \leq A_i, \forall i$ and $B_{i^*} \leq B_i, \forall A_i = A_{i^*}$.

B. Feasibility analysis of concurrent transmissions

In the time slot t, F upcoming links corresponding to F initial paths need to be simultaneously active for data packet transmissions. Accounting for the interference from concurrent transmissions, it is necessary to analyze the feasibility of a given concurrent link set $\mathcal{L}$.

Set $M = |\mathcal{L}|$ as the number of links in $\mathcal{L}$. $\mathcal{L}$ will be feasible only when it satisfies the following two conditions:

Half-duplex Condition. With half-duplex nodes, it is clear that, for any $i, j, m \in \mathcal{V}$, link $\mathcal{P}_{i,j} \in \mathcal{L}$ implies $\mathcal{P}_{j,m} \notin \mathcal{L}$ and vice versa.

Minimum SINR Condition. For any link $\mathcal{P}_{i,j}$, we have

$$\gamma_{i,j} \geq \beta, \qquad (10)$$

where $\gamma_{i,j}$ is the SINR seen by node j for signal from node i.

It is trivial to check whether $\mathcal{L}$ satisfies the half-duplex condition, so in the next we will focus on the minimum SINR condition.

Assuming that $\mathcal{L}$ satisfies the half-duplex condition, since the degree of guard zone $D_{i,j}$ for link $\mathcal{P}_{i,j} \in \mathcal{L}$ is the number of transmitters inside the guard zone $A_{i,j}$, $D_{i,j}$ indicates the level of temporary strong interference on $\mathcal{P}_{i,j}$. If $D_{i,j} = 1$ and there is no links in $\mathcal{L}$ sharing transmitter i with $\mathcal{P}_{i,j}$, node j will directly decode the desired signal. Otherwise, node j will first decode the interfering signal and then cancel the interference and decode the desired signal.

To satisfy (10), each link corresponds to a pair of constraints which can be given according to the degree of guard zone. When the degree of guard zone is larger than 1, the interference signal inside the guard zone will be decoded in the descending order of channel conditions. Suppose that there exists a link $\mathcal{P}_{i,j}$ with $D_{i,j}$, the SINR constraint on each interference signal from transmitter m inside $A_{i,j}$ can be expressed as

$$\frac{P_m \cdot g_{m,j}}{P_i \cdot g_{i,j} + \sum_{k \neq i}^{g_{k,j} < g_{m,j}} P_k \cdot g_{k,j} + \sigma^2} \geq \beta. \qquad (11)$$

With SIC, after $(D_{i,j} - 1)$ interference signals have been successfully decoded, the desired signal can be decoded if the SINR $\gamma_{i,j}$ satisfies

$$\frac{P_i \cdot g_{i,j}}{\sum_{k \notin A_{i,j}} P_k \cdot g_{k,j} + \sigma^2} \geq \beta. \qquad (12)$$

Particularly, in case $\mathcal{P}_{i,j}$ shares a common transmitter with other $(L-1)$ links in $\mathcal{L}$, with SC, transmitter i has to split its power into L parts. Let $\mathcal{J}$ ($|\mathcal{J}| = L$) be the receiver set including j, P_i^k is the allocated power for $\mathcal{P}_{i,k}, k \in \mathcal{J}$, and it has

$$P_i = \sum_{k \in \mathcal{J}} P_i^k. \qquad (13)$$

Receiver j has to decode the interference signals of $\mathcal{J}$ in an ascending order of channel conditions until the desired signal can be successfully decoded. Therefore, the SINR constraint on all the decoded signals at receiver j (including the desired signal) can be given by

$$\frac{P_i^u \cdot g_{i,j}}{\sum_{l \notin A_{i,j}} P_l \cdot g_{l,j} + \sum_{k \in \mathcal{J}}^{g_{i,k} > g_{i,u}} P_i^k \cdot g_{i,j} + \sigma^2} \geq \beta, \qquad (14)$$
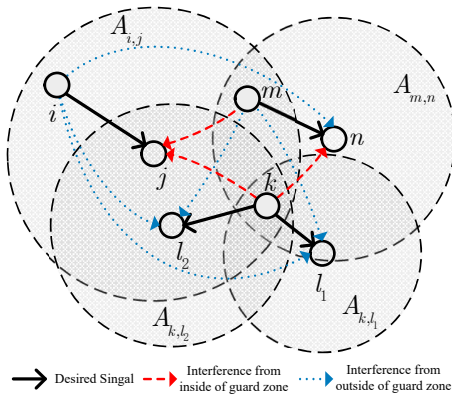
Fig. 2: An example for clarifying the feasibility of concurrent transmission.

where $u \in \{k \in \mathcal{J} \,|\, g_{i,k} < g_{i,j}\}$.

Based on the above analysis, M pairs of the minimum SINR constraints, corresponding to M links in $\mathcal{L}$, can be formulated as M linear inequalities. We can further formulate a LP problem that minimizes the total transmission power subject to M pairs of the minimum SINR constraints. The optimal power allocation for M concurrent transmissions can be obtained by solving the above LP problem. Since each node has a maximum transmission power constraint, $\mathcal{L}$ will be feasible if the transmission power of any transmitter is no more than the maximum transmission power.

Fig. 2 gives an example for clarification of feasibility. It can be seen from Fig. 2 that $\mathcal{L} = \{\mathcal{P}_{i,j}, \mathcal{P}_{m,n}, \mathcal{P}_{k,l_1}, \mathcal{P}_{k,l_2}\}$ satisfies the half-duplex condition. Meanwhile, it can be seen that $D_{i,j} = 3$, $D_{m,n} = 2$, $D_{k,l_1} = 1$ and $D_{k,l_1} = 1$. Therefore, according to (11) and (12), SIC can be used to formulate the SINR constraint for $\mathcal{P}_{m,n}$ as

$$\begin{cases} \dfrac{P_k \cdot g_{k,n}}{P_i \cdot g_{i,n} + P_m \cdot g_{m,n} + \sigma^2} \geq \beta \\ \dfrac{P_m \cdot g_{m,n}}{P_i \cdot g_{i,n} + \sigma^2} \geq \beta. \end{cases} \tag{15}$$

Both nodes m and k are located inside $A_{i,j}$. Therefore, suppose that $g_{m,j} > g_{k,j}$, the SINR constraint for $\mathcal{P}_{i,j}$ using SIC can be formulated as

$$\begin{cases} \dfrac{P_m \cdot g_{m,j}}{P_k \cdot g_{k,j} + P_i \cdot g_{i,j} + \sigma^2} \geq \beta \\ \dfrac{P_k \cdot g_{k,j}}{P_i \cdot g_{i,j} + \sigma^2} \geq \beta \\ \dfrac{P_i \cdot g_{i,j}}{\sigma^2} \geq \beta. \end{cases} \tag{16}$$

$\mathcal{P}_{k,l_1}$ and $\mathcal{P}_{k,l_2}$ share a common transmitter. Therefore, transmitter k will split its power into two parts for receivers l_1 and l_2:

$$P_i = P_k^{l_1} + P_k^{l_2}. \tag{17}$$

Suppose that $g_{k,l_1} > g_{k,l_2}$, according to (14), the SINR

constraint for the link $\mathcal{P}_{k,l_1}$ can be given based on SC as

$$\begin{cases} \dfrac{P_k^{l_2} \cdot g_{k,l_1}}{P_i \cdot g_{i,l_1} + P_m \cdot g_{m,l_1} + P_k^{l_1} \cdot g_{k,l_1} + \sigma^2} \geq \beta \\ \dfrac{P_k^{l_1} \cdot g_{k,l_1}}{P_i \cdot g_{i,l_1} + P_m \cdot g_{m,l_1} + \sigma^2} \geq \beta, \end{cases} \tag{18}$$

and the SINR constraint for $\mathcal{P}_{k,l_2}$ can be given by

$$\dfrac{P_k^{l_2} \cdot g_{k,l_2}}{P_i \cdot g_{i,l_2} + P_m \cdot g_{m,l_2} + P_k^{l_1} \cdot g_{k,l_2} + \sigma^2} \geq \beta. \tag{19}$$

Finally, the transmission power minimization problem can be expressed as a LP problem

$$\begin{aligned} \min \quad & P_i + P_m + P_k \\ s.t. \quad & (15), (16), (17), (18), (19). \end{aligned} \tag{20}$$

Since the number of variables is less than that of inequalities, the above LP problem can be solved by applying the Fourier-Motzkin elimination (FME) method [40] that gradually eliminates variables from the linear inequalities and removes the constraints. The solution to (20) provides the power allocation for the concurrent transmissions of links in $\mathcal{L}$. Only when all of three powers $\{P_i, P_m, P_k\}$ are lower than $P_{\max}$, will $\mathcal{L}$ be deemed a feasible link set.

C. Interference-aware scheduling

Denote $\mathcal{L}_t$ as the candidate link set formed by the upcoming links of F paths at the t-th time slot, and $F = |\mathcal{L}_t|$. Clearly, in order to achieve low-latency end-to-end transmissions, we need to find the maximum non-empty feasible link set from $\mathcal{L}_t$ to perform the interference-aware concurrent transmissions.

A guard zone based selection (GBS) algorithm is introduced to greedily explore the maximum feasible link set $\mathcal{L}'_t$ from $\mathcal{L}_t$. The proposed GBS algorithm is designed based on the prioritization of candidate links. All the links are firstly prioritized by the degree of guard zone obtained from $\mathcal{L}_t$, and a smaller degree of guard zone yields a higher priority. Accounting for that multiple links may have the same degrees, the channel conditions of links are chosen as the second metric for the prioritization, and a weaker channel yields a higher priority. This design is due to the fact that the weak link is vulnerable to the interference and should be scheduled as soon as possible.

The main idea of the proposed GBS algorithm lies in the fact that all the candidate links in $\mathcal{L}_t$ are sequentially selected by priority into $\mathcal{L}'_t$. Every time a candidate link $\mathcal{P}_{\text{tmp}} \in \mathcal{L}_t$ is added, the feasibility of updated $\{\mathcal{L}'_t, \mathcal{P}_{\text{tmp}}\}$ will be checked. If it is feasible, $\mathcal{L}'_t = \{\mathcal{L}'_t, \mathcal{P}_{\text{tmp}}\}$, otherwise, $\mathcal{P}_{\text{tmp}}$ will be put into into $\mathcal{G}'_t$. Finally, the maximum non-empty feasible link set $\mathcal{L}'_t \subset \mathcal{L}_t$ can be obtained. The infeasible links in $\mathcal{G}'_t$ will be put off to the next time slot. Algorithm 2 shows the exact procedure of the proposed GBS algorithm. Clearly, the computational complexity of proposed GBS algorithm is $O(F)$.

Note that the proposed GBS algorithm can be achieved in a distributed fashion to explore the maximum feasible link set and allow the maximum number of concurrent transmissions for each time slot. The detailed distributed design is given as follows.

Algorithm 2: Proposed GBS algorithm

Input: Candidate concurrent link set $\mathcal{L}_t$, $F = |\mathcal{L}_t|$;
**Output:** Feasible link set $\mathcal{L}'_t$, infeasible link set $\mathcal{G}'_t$.

1   Initiate $\mathcal{L}'_t = \emptyset$, $\mathcal{G}'_t = \emptyset$.
2   **for** $m = 1$ *to* F **do**
3     $\mathcal{P}_{\text{tmp}} \leftarrow$ The link with the highest priority in $\mathcal{L}_t$,
4   Remove $\mathcal{P}_{\text{tmp}}$ from $\mathcal{L}_t$ ($\mathcal{L}_t = \{\mathcal{L}_t / \mathcal{P}_{\text{tmp}}\}$).
5   Check the feasibility of $\{\mathcal{L}'_t, \mathcal{P}_{\text{tmp}}\}$.
6   **if** $\{\mathcal{L}'_t, \mathcal{P}_{\text{tmp}}\}$ *is feasible* **then**
7       $\mathcal{L}'_t \leftarrow \{\mathcal{L}'_t, \mathcal{P}_{\text{tmp}}\}$, record power allocation.
8   **else**
9       $\mathcal{G}'_t \leftarrow \{\mathcal{G}'_t, \mathcal{P}_{\text{tmp}}\}$.
10   **end**
11 **end**
12 Obtain the maximum feasible link set $\mathcal{L}'_t$.
13 **if** $\mathcal{G}'_t$ *is not an empty set* **then**
14   Put all the links in $\mathcal{G}'_t$ off to the next time slot.
15 **end**

Before each concurrent transmission, F instant links from F paths will form the candidate concurrent link set by broadcasting the individual link information (transmitter, receiver and channel state information). By resolving the broadcast information, each transmitter i of link $\mathcal{P}_{ij}$ can have the knowledge of the priorities of all the candidate links. Then, the feasibility of the concurrent link set including $\mathcal{P}_{ij}$ and other links with higher priorities will be checked by transmitter i. If the concurrent link set is feasible and no transmitter with lower priority has explicitly broadcast the feasible results, transmitter i will broadcast its feasible results that contains the exact power allocation. Finally, together with transmitter i, all the transmitters with the higher priority will perform the concurrent transmissions with the corresponding power allocation. The remaining links will be put off to the next time slot. Repeatedly, by running the proposed GBS algorithm slot by slot, the interference-aware scheduling for multiple paths can be achieved in a compact manner.

IV. SIMULATION RESULTS

In this section, the simulations are designed to show the effectiveness of the proposed scheme by comparing its performance with the existing schemes via simulation. In the simulation, FME is performed by using MATLAB FME-IT package [41]. We simulate a randomly generated multi-hop network with N nodes, which are distributed in a $100 \times 100 \text{m}^2$ square area (from the origin to (100,100)). Quoted or inferred from [32]–[38], the related system parameters are set as: the path-loss exponent $\alpha = 3.5$, the SINR threshold $\beta = 2(3\text{dB})$, the power of ambient noise $\sigma^2 = -60\text{dBm}$, and the maximum transmission power of each node is set to $P_{\text{max}} = 20\text{dBm}$.

The time is divided into slots, and each slot represents one second in the simulations. Considering that each transmission costs a time slot, the hop count of initial paths can indicate the lower bound of required end-to-end delay for each data session. Therefore, on the basis of initial paths constructed by interference-aware routing algorithm, we deploy the link
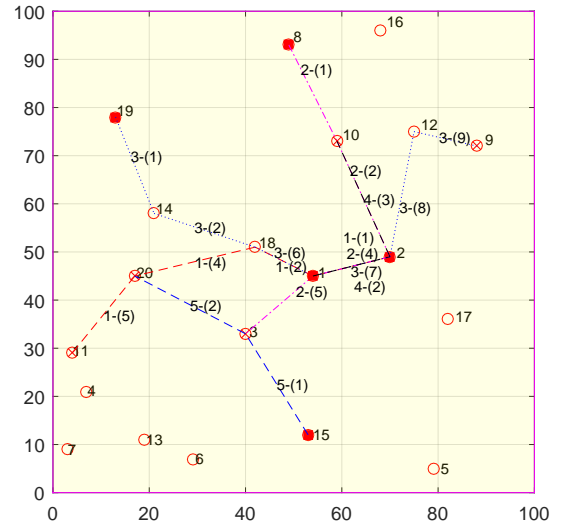
(a) Overall scheduling result.
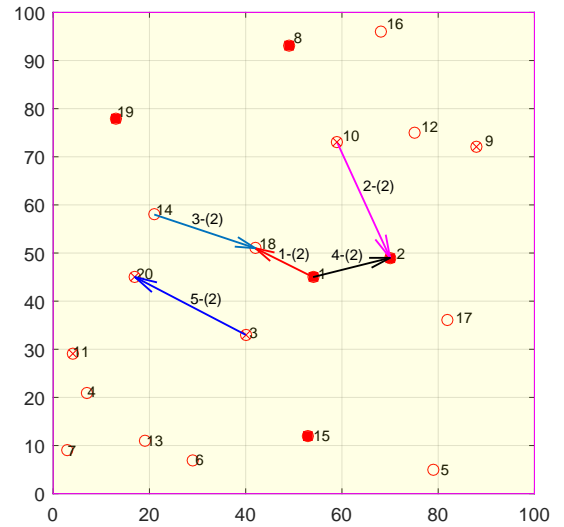
(b) Detailed active links in the 2-nd time slot.

Fig. 3: A five-data-packet routing and scheduling result for the 20-node network under the proposed ICR scheme.

scheduling in the MAC layer with only the half-duplex constraints satisfied as the Lower Bound of performance. Meanwhile, the performance of traditional SAR scheme [36] is evaluated and compared with our proposed ICR scheme. Furthermore, we adopt the interference-aware routing with link scheduling (IR-SCH) and spatial resource-aware routing with link scheduling (SR-SCH) schemes as the baselines. These two schemes firstly construct the initial end-to-end path by the proposed interference-aware routing algorithm and the spital resource-aware routing algorithm (which is used to construct the initial path in SAR scheme), respectively, and then the traditional scheduling [34], where both SIC and SC are not employed, is deployed to handle all the interference in the network.

TABLE I: End-to-end delay for five given source-destination packets (Time slots)

Parameter	Data 1	Data 2	Data 3	Data 4	Data 5
$\left\{V_s^{(f)}, V_d^{(f)}\right\}$	{2, 11}	{8, 3}	{19, 9}	{1, 10}	{15, 20}
Lower Bound	4	4	6	3	2
ICR	**5**	**5**	**9**	**3**	**2**
SAR	6	6	10	3	2
IR-SCH	6	7	11	6	3
SR-SCH	6	7	11	8	4

A. An illustrative example

We first use a 20-node network depicted in Fig. 3 as an example to illustrate the scheduling results and the advantages in reducing the end-to-end delay of the proposed ICR scheme.

Here we see five concurrent source-destination pairs $\left\{V_s^{(f)}, V_d^{(f)}\right\}$, $f = 1, \cdots, 5$ in Table I, and each source has one data packet to transmit to the destination. Fig. 3a shows the overall scheduling result and Fig. 3b shows the detailed active links in the 2-nd time slot. In Fig. 3a, the red solid points represent the source nodes, and the red hollow points with '×' labeled are the destination nodes. Different source-destination packets are shown in different color and the corresponding packet ID plus time slot scheduled are labeled on each active link. It can be shown from the Fig. 3b that in the 2-nd time slot, the link pairs $\{\mathcal{P}_{1,18}, \mathcal{P}_{14,18}\}$ and $\{\mathcal{P}_{10,2}, \mathcal{P}_{1,2}\}$ share the common receiver 18 and 2, respectively. $\{\mathcal{P}_{1,18}, \mathcal{P}_{1,2}\}$ shares the common transmitter 1. In these cases, SIC and SC can increase spatial reuse and save spatial resource so that the end-to-end delay is significantly reduced.

Furthermore, the exact end-to-end delay for five data packets under different routing schemes is shown in Table I. It can be seen that our proposed ICR scheme outperforms the remaining schemes except the Lower Bound. The performance gap between Lower Bound and ICR scheme is due to the interference-aware design. Moreover, by comparing SR-SCH with IR-SCH, we can find that the construction of initial paths in SAR scheme increases the transmission delay, while in contrast to the SAR scheme, it can be proven that the SIC-aware consideration has also achieved an effective latency reduction.

B. Performance evaluation

Next, the traffic model in the network is set as the poisson traffic model with the mean packet arrival rate being λ, and the packet generation process for each node per time unit follows a poisson distribution with mean arrival rate being λ/N. Suppose that all the packets have an identical destination node N which is fixed at the right-up corner (100,100), and the remaining $N-1$ nodes are randomly distributed in the square area. All the outputs are based on 50 randomly generated topologies each with 100 trials, and in each trail, the packet generation lasts for 40 seconds.

A snapshot captured from one of the simulation topologies and trails with $N = 80$ and $\lambda = 1$ is given in Fig. 4. We can observe that the right-up region of the network is the hot spot area because all the packets will finally get to the destination node 80.
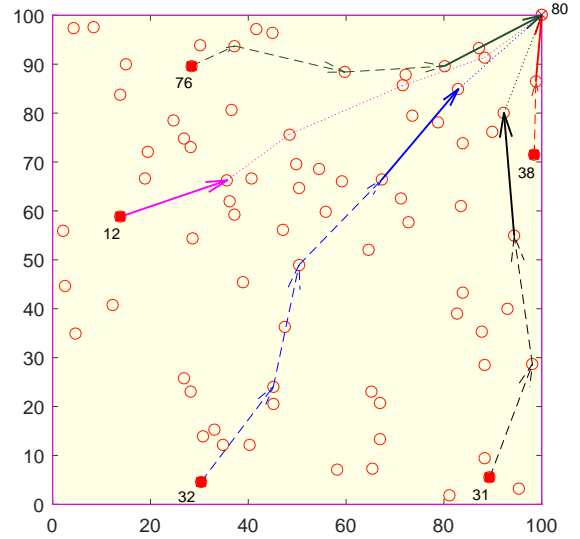

Fig. 4: A snapshot captured from one of the simulation topologies and trails with $N = 80$ and $\lambda = 1$. The dotted lines with an arrow represent the past links originated from the sources, the solid lines with an arrow are the active links in the current time slot, and the dotted lines without an arrow are the waiting transmissions for the paths.

1) Average end-to-end delay: Fig. 5 depicts the average end-to-end delay versus the different number of nodes N ranging from 20 to 140, with the mean packet arrival ratio $\lambda = 1$. It can be seen from Fig. 5 that the curves of five schemes are decreasing as N increases, this is because when the density of nodes increases, the path with less hop count can be found which results in the lower transmission delay. Moreover, Fig. 5 also shows that the end-to-end delay with the proposed ICR scheme is the lowest compared with the other schemes except the Lower Bound, which verifies the effectiveness of the proposed scheme in lowering the end-to-end delay. Statistically, our ICR scheme lowers the end-to-end delay from 9.16% to 73.82%, compared with the other schemes. Noticeably, the ratio raises with the increasing number of nodes. The ICR scheme has achieved a lower end-to-end delay than the IR-SCH scheme because of the advantages of SIC and SC. Meanwhile, the ICR scheme outperforms the SAR scheme due to the effectiveness of the proposed GBS algorithm. Furthermore, the IR-SCH scheme outperforms the SR-SCR scheme which verifies the effectiveness of the proposed interference-aware routing.

Fig. 6 shows the impact of the different mean packet arrival ratio λ on the average end-to-end delay for the different routing schemes. λ varies from 0.25 to 2, with the number of nodes $N = 80$. We can observe from Fig. 6 that the end-to-end delay increases with a larger value of λ, this is because the increase of data traffic will bring more interference to the current links, so that the links in the hot spot regions (generally in the region close to destination node) will be delayed more for the heavier traffic load. Moreover, the proposed ICR scheme has achieved the smallest performance gap with the Lower Bound. Compared with the SAR, IR-SCH and SR-
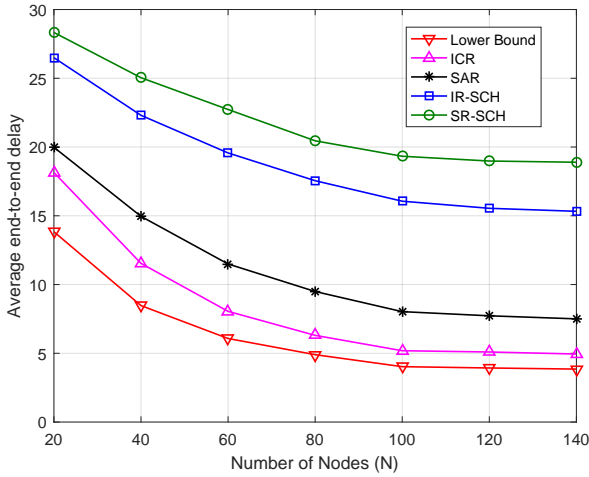
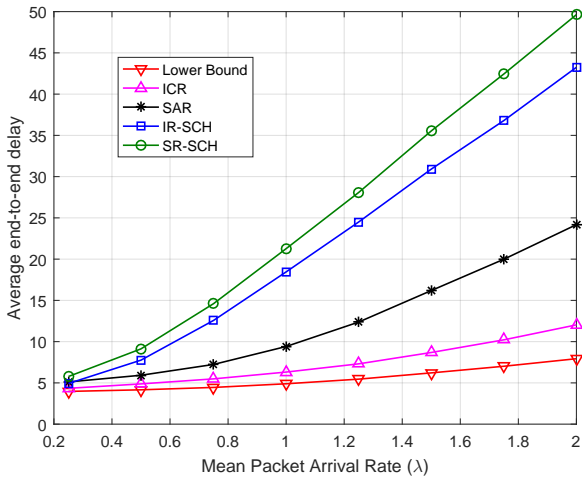Fig. 5: Average end-to-end delay versus the number of nodes with $\lambda = 1$.

Fig. 7: The maximum end-to-end delay versus mean packet arrival ratio λ with $N = 80$.

Fig. 6: Average end-to-end delay versus mean packet arrival ratio with $N = 80$.

the maximum end-to-end delay during a trail, the cumulative distribution function (CDF) of end-to-end delay and the packet delivery ratio are evaluated via simulations.

Fig. 7 shows the maximum end-to-end delay versus varied λ. It can be observed that the maximum end-to-end delay increases with more data traffic. The performance gap between ICR and Lower Bound reflects the impacts of interference. Meanwhile, the performance gap between IR-SCH and SR-SCH verifies the gains of proposed interference-aware routing. More importantly, apart from the Lower Bound, the proposed ICR scheme has achieved the lowest delay among the four interference-aware schemes, which proves its best capability of concurrent transmissions during one time slot and excellent low-latency performance in the multi-hop networks.

To clearly show the performance difference of five routing schemes, we plot the simulated CDF of end-to-end delay with $N = 80$ and $\lambda = 1$ in Fig. 8. Note that the curve labeled by Lower Bound represents the imaginary scheduling scheme to reflect the lower bound of end-to-end delay. As can be seen, the CDF of performance with the proposed ICR scheme has achieved the highest similarity with that of Lower Bound. Moreover, in the network with the proposed ICR scheme, more than 90% of the packets can be successfully received by the destination node within 10 time slots. In contrast, the delay constraints in the network with the SAR, IR-SCH and SR-SCH schemes need to be extended to 19, 47 and 51 time slots, which shows the effectiveness of the ICR scheme in reducing the end-to-end delay.

3) Packet delivery ratio: The packet delivery ratio under a given maximum time constraint T is further evaluated with different λ. A packet can be seen as a successful delivery, only if the packet is successfully received by the destination node within T time slots.

Fig. 9 shows the simulation results with $N = 80$ and $T = 30$. It can be seen from the figure that the larger λ results in more interference, the end-to-end delay becomes longer and thus leads to a lower packet delivery ratio. The network with the proposed ICR scheme can support the largest traffic

SCH schemes, our proposed ICR scheme can reduce end-to-end delay by 32.93%, 65.80% and 70.33%, respectively, for $\lambda = 1$, and the reduction is 50.24%, 72.18% and 75.76%, respectively, for $\lambda = 2$, the reason for which lies in that, on the one hand, compared with the IR-SCH, the interference coordination allows multiple nodes to share a common transmitter or receiver which can save the spatial resource, so that more data packets concurrence can be supported. On the other hand, the proposed GBS algorithm explores the maximum feasible link set for each time slot which also improves the capability of concurrent transmissions during one time slot.

2) Distribution of end-to-end delay: Considering that the reliable and timely delivery of data packet for some mission-critical applications, e.g., target tracking and emergency alarm is crucial in the success of the mission, the maximum end-to-end delay constraint usually exists in the network. In this case, it is significant to evaluate these routing schemes by analyzing the exact distribution of end-to-end delay. Therefore, to further show the effectiveness of the proposed ICR scheme,
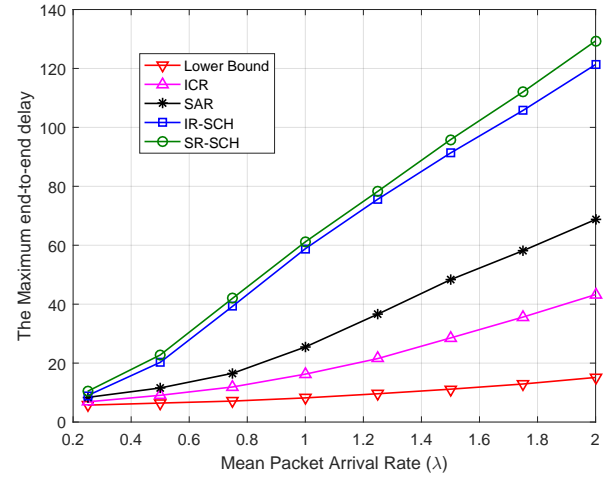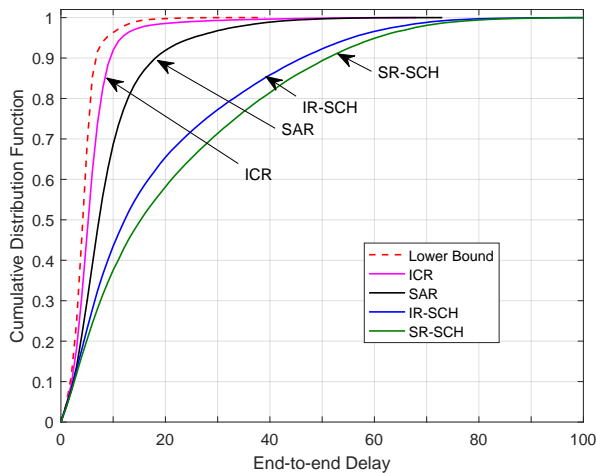
Fig. 8: The cumulative distribution function versus the end-to-end delay with $\lambda = 1$ and $N = 80$.
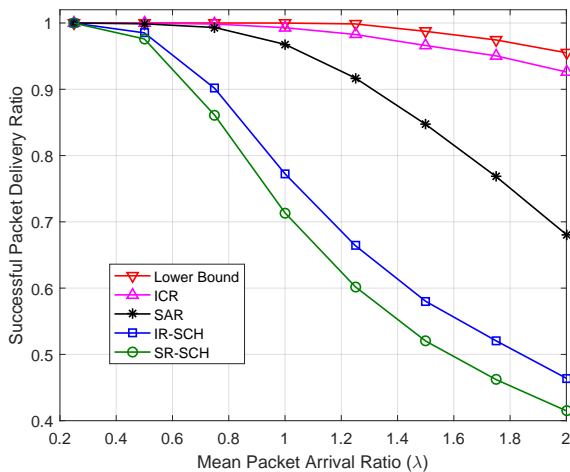

Fig. 9: The successful packet delivery ratio versus the mean packet arrival ratio λ with time constraint $T = 30$ and $N = 80$.

under a fixed time constraint, which reflects the excellent capability of concurrent transmissions for the multi-hop scenarios. Meanwhile, the ICR scheme outperforms the traditional SAR scheme because the proposed GBS algorithm is more effective in handling the interference coordination, and the packet delivery ratio for the network with the IR-SCH scheme is higher than the SR-SCH scheme proves the effectiveness of proposed interference-aware routing. Therefore, it can be concluded that the proposed ICR scheme can significantly improve the end-to-end reliability and achieve a low-latency transmission.

V. CONCLUSION

In this paper, we presented a distributed ICR scheme to reduce the end-to-end delay in wireless multi-hop networks by combining interference-aware routing in network layer, link scheduling in MAC layer, and power allocation in physical layer. The proposed ICR scheme firstly introduced an interference-aware routing algorithm that captures the end-to-end latency and spatial resource cost as the routing metrics to construct the end-to-end initial path. Then, by allowing that each node in the network is endowed with interference cancelation capabilities, the concurrent transmission of multiple active links from different paths was formulated as a LP problem, and the solution to the LP problem indicated the power allocation for each link. Finally, a distributed GBS algorithm was further proposed to iteratively explore the maximum feasible link set for each time slot. The selected links will be active for transmission according to the allocated power in the current time slot, and the packets of the remaining links will be put off to the next time slot. The simulation results verifies that the proposed ICR scheme reduces the end-to-end delay by at least 9.16% and up to 73.82 when λ is 1 compared with the existing schemes. We have also shown that networks with the proposed ICR scheme promote better transmission concurrence and, therefore, better spatial reuse.

It is worth mentioning that our proposed interference-aware routing considers the effect of interference from the perspective of possible active links, especially the delay caused by the incompatibility among links due to the half-duplex condition. In fact, only the power allocation could truly reflect the feasibility of concurrent transmission for multiple links. To this end, our future work will design the routing protocol jointly considering the routing selection, link scheduling, and power allocation.

REFERENCES

[1] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How Can Heterogeneous Internet of Things Build Our Future: A Survey," *IEEE Communication Survey & Tutorials*, vol. 20, no. 3, pp. 2011-2027, 2018.

[2] J. G. Andrews *et al.*, "What will 5G Be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, 2014.

[3] S. Huang, A. Liu, S. Zhang, T. Wang, and N. Xiong. "BD-VTE: A Novel Baseline Data based Verifiable Trust Evaluation Scheme for Smart Network Systems," *IEEE Trans. on Network Science and Engineering*, DoI: 10.1109/TNSE.2020.3014455, 2020.

[4] B. Jiang, G. Huang, T. Wang, J. Gui, and X. Zhu. "Trust based energy efficient data collection with unmanned aerial vehicle in edge network," *Trans. on Emerging Telecomm. Technologies*, issue e3942, March 2020.

[5] J. Tan, et al.. "A High-accurate Content Popularity Prediction Computational Modelling for Mobile Edge Computing by Using Matrix Completion Technology," *Trans. on Emerging Telecomm. Technologies*. issue e3871, Jan. 2020.

[6] F. S. Shaikh, R. Wismuller, Routing in Multi-Hop Cellular Device-to-Device (D2D) Networks: A Survey. *IEEE Communication Survey & Tutorials*. vol. 20, no. 4, pp. 2622-2657, 2018.

[7] S. Taneja, A. Kush, "A Survey of Routing Protocols in Mobile Ad Hoc Networks," *Int. J. of Innovation, Management and Technology*, Vol. 1, No. 3, pp. 279-285, August 2010.

[8] M. Dehghan, M. Ghaderi, and D.L. Goeckel, "On the performance of Cooperative Routing in Wireless Networks," in *2010 Proc. of IEEE INFOCOM*, March 2010.

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

[10] F. Mansourkiaie, M. H. Ahmed, "Cooperative Routing in Wireless Networks: A Comprehensive Survey," *IEEE Commun. Survey & Tutorials*, Vol. 17, No. 2, pp. 604-626, 2015.

[11] G. Parissidis, M. Karaliopoulos and T. Spyropoulos, "Interference Aware Routing in Wireless Multihop Networks," *IEEE Trans. Mobile Computing*, vol. 10, no. 5, pp. 716-733, May 2011.

[12] M. X. Cheng and X. Gong, "Interference-Aware Multipath Routing and Link Rate Control in Multihop Wireless Networks," in *2010 Proc. IEEE Globecom*, 2010.

[13] R. Langar, N. Bouabdallah and R. Boutaba, "Interferer Link-Aware Routing in Wireless Mesh Networks," in *2010 Proc. IEEE ICC*, 2010.

[14] F. Mansourkiaie and M. H. Ahmed, "Optimal and Near-Optimal Cooperative Routing and Power Allocation for Collision Minimization in Wireless Sensor Networks," *IEEE Sensor Journal*, vol. 16, no. 5, pp. 1398-1411, March 2016.

[15] Y. Hao, Q. Ni, H. Li, S. Hou, and G. Min, "Interference-Aware Resource Optimization for Device-to-Device Communications in 5G Networks," *IEEE Access*, vol. 6, 2018.

[16] F. Mansourkiaie and M. H. Ahmed, "Joint cooperative routing and power allocation for collision minimization in wireless sensor networks with multiple flows," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 6-9, Feb. 2015.

[17] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1452-1463, Aug. 2006.

[18] J. Zuo, C. Dong, S. X. Ng, L. L. Yang, and L. Hanzo, "Cross-Layer Aided Energy-Efficient Routing Design for Ad Hoc Networks," *IEEE Commun. Survey & Tutorials*, Vol. 17, No. 3, pp. 1214-1238, 2015.

[19] R. I. Ansari *et al.*, "5G D2D Networks: Techniques, Challenges, and Future Prospect," *IEEE System Journal*, vol. 12, no. 4, pp. 3970-3984, Dec. 2018.

[20] D. Gong and Y. Yang, "Low-Latency SINR-Based Data Gathering in Wireless Sensor Networks," *IEEE Trans. Wireless Communications*, vol. 13, no. 6, pp. 3207-3221, June 2014.

[21] B. Han, J. Li, J. Su, and J. Cao, "Self-supported cooperative networking for emergency services in multi-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 450-457, Feb. 2012.

[22] M. Andrews, M. Dinitz, "Maximizing Capacity in Arbitrary Wireless Networks in the SINR Model: Complexity and Game Theory," in *2009 Proc. IEEE INFOCOM*, 2009, pp. 1332-1340.

[23] A. Capone, C. Lei, S. Gualandi, and Y. Di, "A new computational approach for maximum link activation in wireless networks under the SINR model," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1368-1372, 2011.

[24] C. Jiang *et al.*, "Squeezing the most out of interference: An optimization framework for joint interference exploitation and avoidance," in *2012 Proc. IEEE INFOCOM*, 2012, pp. 424-432.

[25] Q. He, D. Yuan, A. Ephremides, "Maximum link activation with cooperative transmission and interference cancellation in wireless networks," *IEEE Trans. Mob. Comput.*, vol. 16, no. 2, pp. 408-421, Feb. 2017.

[26] J. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 19-29, Apr. 2005.

[27] X. Su and S. Chan, "High-Throughput Routing with Superposition Coding and Successive Interference Cancellation," in *2011 Proc. IEEE ICC*, June 2011, pp. 1-6.

[28] S. Lv *et al.*, "Understanding the Scheduling Performance in Wireless Networks with Successive Interference Cancellation," *IEEE Trans. Mobile Computing*, vol. 12, no. 8, pp. 1625-1639, August 2013.

[29] S. Lv, W. Zhuang, X. Wang, and X. Zhou, "Scheduling in Wireless Ad Hoc Networks with Successive Interference Cancellation," in *2011 Proc. IEEE INFOCOM*, 2011, pp. 1287-1295.

[30] S. Lv, W. Zhuang, X. Wang, and X. Zhou, "Context-aware Scheduling in Wireless Networks with Successive Interference Cancellation," in *2011 Proc. IEEE ICC*, 2011, pp. 1-5.

[31] A. Capone, Y. Li, M. Pioro, and D. Yuan, "Minimizing end-to-end delay in multi-hop wireless networks with optimized transmission scheduling," *Ad Hoc Networks*, vol. 89, pp. 236-248, Feb. 2019.

[32] X. Li, M. Sheng, X. Wang, and J. Liu, "Joint Scheduling and Power Control for α-Utility Maximization in Wireless Ad-hoc Networks with Successive Interference Cancellation," in *2014 Proc. IEEE WCNC*, 2014, pp. 1568-1572.

[33] D. Yuan, V. Angelakis, L. Chen, E. Karipidis, and E. Larsson, "On optimal link activation with interference cancelation in wireless networking," *IEEE Tran. Veh. Technol.*, vol. 62, no. 2, pp. 939-945, Feb. 2013.

[34] C. Jiang *et al.*, "Cross-layer optimization for multi-hop wireless networks with successive interference cancellation," *IEEE Tran. Wireless Communications*, vol. 15, no. 8, pp. 5819-5831, Aug. 2016.

[35] L. Qu, J. He, C. Assi, "Congestion Control, Routing, and Scheduling in Wireless Networks With Interference Cancelation Capabilities," *IEEE Tran. Veh. Technol.*, vol. 64, no. 7, pp. 3108-3119, July 2015.

[36] R. Liu *et al.*, "Bandwidth-Aware High-Throughput Routing With Successive Interference Cancelation in Multihop Wireless Networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5866-5877, Dec. 2015.

[37] J. Cheng, Y. Gao, N. Zhang, and H. Yang, "An Energy-Efficient Two-Stage Cooperative Routing Scheme in Wireless Multi-Hop Networks," *Sensors*, vol. 19, no. 5, pp. 1-15, Apr. 2019.

[38] Andrea Goldsmith, "Wireless Communiations," *Standford University, Cambridge University Press*, pp. 27–55, 2005.

[39] C. E. Perkins, E. M. Royer, "Ad-hoc on-demand distance vector (AODV) routing," in *Proc. of 1999 IEEE Workshop on Mobile Computing Systems and Applications*, Feb. 1999.

[40] A. Schrijver, "Theory of linear and integer programming," *John Wiley & Sons*, pp. 155–156, 1998.

[41] I. B. Gattegno, Z. Goldfeld, and H. H. Permuter, "Fourier-Motzkin Elimination Software for Information Theoretic Inequalities," *IEEE Inf. Theory Soc. Newslett.*, vol. 65, no. 3, pp. 25–28, Sep. 2015.
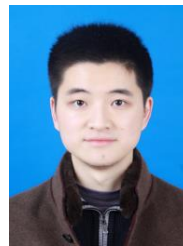
Jianming Cheng received his M.E. degree in Electronic and Communication Engineering at Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2016. He is currently working forwards the Ph.D. with the school of Wireless Communication Center, BUPT, Beijing, China. He has been a Visiting Ph.D. Student with the School of Computing and Communications, Lancaster University, Lancaster, U.K., since 2019. His research interests include wireless ad-hoc networks, cross-layer design, and green communications.

Pei Yang received his B.S. degree in Information and Communication Engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2016. He is currently pursuing a Ph.D. degree with the school of Wireless Communication Center, BUPT, Beijing, China. He has been a visiting Ph.D. student in the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA, since 2018. His research interests lie in wireless communications, information theory, and signal processing.

Keivan Navaie is with the School of Computing and Communications, Lancaster University, UK. His research interests include Distributed Cloud and Edge Computing, Connectivity Resilience in Cyber-Physical Systems and IoT, Applications of Machine Learning and Artificial Intelligence, and Cognitive Communications Networks. He is a *Fellow of IET, Senior Member of the IEEE, Senior Fellow of HEA*, and a *Chartered Engineer* in the UK. He currently serves on the editorial board of the IEEE Comm. Letters, IEEE Comm. Surveys and Tutorials, Trans. on Wireless Communications, and IEEE Access. He is also active in evidence-based policy making through his involvement with the EU Environment Agency and EU Parliament as an Independent Scientist Expert.

Qiang Ni (M'04–SM'08) received the B.Sc., M.Sc., and Ph.D. degrees from the Huazhong University of Science and Technology, China, all in engineering. He is currently a Professor and the Head of the Communication Systems Group, School of Computing and Communications, Lancaster University, Lancaster, U.K. His research interests include the area of future generation communications and networking, including green communications and networking, millimeter-wave wireless communications, cognitive radio network systems, non-orthogonal multiple access (NOMA), heterogeneous networks, 5G and 6G, SDN, cloud networks, energy harvesting, wireless information and power transfer, IoTs, cyber physical systems, AI and machine learning, big data analytics, and vehicular networks. He has authored or co-authored over 200 papers in these areas. He was an IEEE 802.11 Wireless Standard Working Group Voting Member and a contributor to the IEEE Wireless Standards.

Hongwen Yang received the B.S. and M.S. degrees from the Beijing University of Posts and Telecommunications (BUPT), in 1984 and 1987, respectively. After graduating, he joined the faculty of BUPT, where he is currently a Professor. His research mainly focuses on the wireless physical layer, including modulation and coding, MIMO, and OFDM.