Feasibility of Multipath Construction in mmWave Backhaul

Yan Yan, Qiang Hu and Douglas M. Blough
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332

Abstract—This paper focuses on the problem of finding multiple paths with relay nodes to maximize throughput for ultra-high-rate millimeter wave (mmWave) backhaul networks in urban environments. Relays are selected between a pair of source and destination base stations to form multiple interference-free paths. We first formulate the problem of feasibility of multi-path construction as a constraint satisfaction problem that includes constraints on intra-path and inter-path interference and several other constraints that arise from the problem setting. Based on the derived equations, we transform the multiple paths construction problem into a Boolean satisfiability problem. This problem can then be solved through use of a satisfiability (SAT) solver, which however results in a very high running time for realistic problem sizes. To address this, we propose a heuristic algorithm that runs in a fraction of the time of the SAT solver and finds multiple interference-free paths using a modification of a maximum flow algorithm. Simulation results based on 3-D models of a section of downtown Atlanta show that the heuristic algorithm finds multiple paths in almost all the feasible cases (those where the SAT solver succeeds in finding a solution) and produces paths with higher average throughput than the SAT solver. Furthermore, the heuristic increases throughput by 50-100% in typical cases compared to a single-path solution.

I. INTRODUCTION

Globally, the demand for higher network capacity is continually increasing. The next generation ultra-high bandwidth communication system, known as 5G, is planned to increase the capacity of existing networks by 1000 fold within the next 20 years [1]. 5G mobile communications will support a wide range of use cases and emerging applications including augmented reality, data sharing, machine-to-machine applications and real-time HD video streaming [2]. The major capacity-enhancing tools that are projected to meet 5G data traffic requirements include a migration to the millimeter-wave (mmWave) bands, which have a large available bandwidth, and a dense infrastructure deployment, where cell size is substantially decreased as compared to 4G networks.

Due to the deployment flexibility, lower maintenance costs and the ability to improve network capacity, small cells are being widely considered and used by operators [3]. However, despite their benefits, one of the main challenges in densely deploying small-cell base stations (SBSs) is what type of backhaul network should be deployed to carry traffic between SBSs and the core network. Deploying fiber backhaul links to every SBS is costly and difficult and might not even be possible in some scenarios, e.g. in older urban areas without advanced infrastructure support. Thus, alternative backhaul architectures will be required for many 5G deployments.

To address the 5G backhaul problem, the use of mmWave transmission, in the 30–300 GHz range, has been proposed [4], [5]. mmWave communication can provide higher data rates than is possible at lower frequencies such as the currently used sub 6 GHz bands. For this reason, mmWave communication is considered an ideal technology for 5G backhaul connection for dense small cells. However, there are some fundamental challenges with mmWave communication, such as high propagation loss, directionality, and susceptibility to blockage of mmWave signals. These challenges make mmWave communication best suited for short-range transmissions. To facilitate mmWave communications over longer distances, a number of researchers have proposed the use of relay nodes [6], [7].

Relay nodes have a simple structure, are easily deployable, and have lower cost compared with base stations (BSs). In recent years, using relays in mmWave backhaul has attracted significant attention [7], [8]. Instead of a single direct long transmission from a transmitter to a receiver, relay nodes can be used to send the source information to the destination via a sequence of shorter communication hops. These shorter hops can avoid obstacles in the dense urban environment in order to maintain line-of-sight (LoS) communication on each hop, which is required to achieve the ultra high rates necessary for mmWave backhaul networks. The scenario considered herein is 3D outdoor long range multi-hop mmWave communication. Previous work on relay path construction for mmWave backhaul focused on construction of a single optimal path between a base station pair [9], [10], whereas this paper focuses on construction of multiple paths between each pair. Therefore, this work must account for additional constraints between paths not considered in prior work.

Fig. 1. Multipath Construction in Relay Assisted mmWave Backhaul Network

Figure 1 illustrates a localized piece of the reference scenario considered in this paper. SBSs are densely deployed in an urban area. Only some BSs have fiber connections to the broader network and, therefore, other BSs need to communicate wirelessly to send traffic to/from the wired...
connectivity points. Such an architecture is referred to as a self-backhaul network [11]. To facilitate line-of-sight communications between BSs in this dense urban environment, relay nodes are deployed at strategic points. As mentioned, prior work has considered how to select the best relay locations to support a single high-data-rate path between a pair of SBSs using line-of-sight communication at each hop. However, such a solution might still not be able to support the ultra-high data rates necessary for backhaul traffic, particularly for the links close to the wired connectivity points, where traffic is typically aggregated from a fairly high number of SBSs.

In this paper, we explore the possibility of deploying multiple relay paths between a single pair of SBSs in this wireless backhaul scenario. Multiple independent and interference-free paths can be used in parallel, potentially doubling or tripling the achievable data rate between a pair of BSs. Figure 1 depicts an example, where 3 independent relay paths are deployed between two SBSs. Subsequent sections present the formulation of the multi-path problem as a Boolean satisfiability problem, a heuristic algorithm for finding multiple constraint-satisfying paths, and extensive simulation results based on a realistic 3-D urban topology. The specific contributions of this paper include:

- the first exact formulation of the multi-path construction problem in mmWave backhaul networks with constraints,
- a reformulation of the multi-path problem as a Boolean satisfiability problem, which allows satisfiability (SAT) solvers to check for existence of multiple paths that satisfy the constraints,
- design of a heuristic algorithm to efficiently find constraint-satisfying multiple paths in many cases, and
- extensive simulation results, which demonstrate that:
  - the SAT solver can judge the existence of multiple paths precisely, but it can take more than one hour to produce a result for some of the cases we evaluated,
  - the heuristic multi-path construction algorithm produces substantial throughput improvements for mmWave backhaul connections, and
  - the heuristic algorithm is able to find multiple paths satisfying the given constraints in almost all cases where they exist and it produces its result within a few seconds for all evaluated cases.

II. RELATED WORK

In recent years, mmWave backhaul networks for small cell deployments have received significant attention. Two network architectures have been proposed for this scenario: centralized and distributed. In the centralized network architecture, a macro BS is located in the center with small-cell BSs located around it. Direct links between two small-cell BSs are not allowed and, therefore, backhaul data from each small-cell BS is sent directly to the macro BS via a wireless link. In distributed network architectures, backhaul data can be transmitted between nearby small-cell BSs through wireless links and reach the designated BS possibly via multi-hop wireless paths. In [12], system-level simulations showed that distributed network architectures achieve higher throughput gain and are more flexible than the centralized architecture. Therefore, we adopt the distributed architecture herein.

In [13]–[15], robust transmission of mmWave are considered for outdoor roadside and indoor deployments. However, the dense urban deployments considered herein are significantly different than roadside outdoor and indoor deployments due to the abundant obstacles blocking LoS paths in dense urban environments. In [16], a point-to-multipoint in-band mmWave backhaul for 5G networks is proposed and a time-division multiplexing based scheduling scheme is proposed for inter-BS communications. However, the directivity of mmWave and the interference between concurrent transmission are not considered. In [17], mmWave backhaul network optimization models are used to minimize the number of wired BSs and maximize the network flow. In [18], the mmWave backhaul scheduling problem is shown to be NP-hard and heuristic algorithms are evaluated through simulation. In both [17] and [18], the term “relaying” is used but it refers to the BSs acting as relay nodes, which is different from our work, where we consider dedicated mmWave relay devices deployed in the backhaul network. In [19], they studied the mobility based wireless systems to boost indoor network performance. The closest works to ours are the aforementioned [9], [10], where constructions of optimal single paths are presented.

The specific problem considered herein is how to construct multiple paths between a given pair of BSs with some constraints, e.g. that different paths cannot interfere with each other and each relay node serves only one path. We formulate this problem into a Boolean satisfiability problem with these and other constraints based on our target urban backhaul scenario and solve it using modern SAT-solvers. A SAT solver is also used to find multiple paths in [20], [21]. However, in [20], [21], the SAT method is used to search the multipath delays in a spread spectrum system, which is quite far from the problem considered herein. To solve the problem in short running time, we also develop a heuristic multiple path construction algorithm, based on a modified maximum flow approach. We assess the capabilities of these approaches in terms of throughput and feasibility of constructing multiple paths. The maximum flow algorithm we use in our heuristic is used in a similar manner in [22]. However, their goals are to find the capacity limits in a network and maximize packet delivery ratio, whereas we use maximum flow to identify multiple paths between a source and destination.

There is similarity between our problem, which is to construct multiple paths by choosing a set of relay locations out of many possible relay locations, and the well-studied problem of multipath routing. For example, in [23], a routing algorithm that produces multiple network-level paths to fulfill the delay requirements in future 5G scenarios is presented. However, in [23], there is only one path between each pair of source-destination BSs, in contrast to our multiple path per BS pair problem. In [24], [25], multiple paths are considered but transmissions on these paths occur at different times, meaning
that interference between paths is not a consideration. Several studies [26], [27] investigate the multipath routing problem in mobile ad hoc networks with directional antennas. However, these works employ distributed flooding-based approaches that would take exponential time to emulate in our centralized setting and, due to their distributed path discovery process, they cannot guarantee to eliminate inter-path interference, which is a primary consideration in our problem formulation.

III. PRELIMINARIES

A. Network model and environment

We consider the scenario where SBSs are densely deployed in an urban area and millimeter wave (mmWave) wireless backhaul is used to avoid costly and difficult to deploy fiber backhaul. With mobile applications requiring ever-increasing amounts of data, the backhaul network must deliver very high throughputs. As stated earlier, relay nodes are a promising approach to provide short LoS links that can sustain the required high throughputs. Even so, conventional single path transmission may not be able to satisfy the ultra-high demands. Therefore, it is interesting to study multi-path delivery since parallelism of packet transmission can further increase network throughput.

Since mmWave signals in 5G scenarios are highly directional and nodes can be placed at different heights in urban settings, we consider the three-dimensional (3D) effects in our system model. We build a 3D topology of buildings in downtown Atlanta, Georgia, and use it in the simulations presented later. The topology covers an area 1200m × 1600m, which includes 227 buildings higher than 5 meters. All of these buildings are modeled as cuboids for simplicity. For better cellular coverage, for each building with a height between 20 and 200 meters, one of its rooftop corners is randomly picked as a candidate location for deploying a BS and the diagonal corners of these rooftops are picked as possible relay locations (183 positions in total). Considering the high path loss of mmWave signals, if a building is higher than \( h_t \) (e.g., \( h_t = 50 \) m), its candidate BS location will be set at the height \( h_L \) (e.g., \( h_L = 50 \) 1600m), and each grid has only one height \( h_L \) (e.g., \( h_L = 50 \) 1600m), and each grid has only one BS selected randomly. Therefore, we have 42 BS positions in total.

![Macro-cell BS](image)

![Small-cell BS](image)

![mmWave Relays](image)

Fig. 2. Building topology in downtown Atlanta.

B. Channel model and propagation assumptions

Our channel model follows earlier work [9] [10], which facilitates comparisons of single path and multiple path solutions. We use the standard assumption of additive white Gaussian noise. Link capacities are assumed to follow Shannon’s Theorem, i.e.

\[
C = B \log_2(1 + \min \{\text{SINR}, \text{SINR}_{\text{max}}\})
\]

where \( B \) is the bandwidth of the channel and \( \text{SINR} \) is the signal to interference plus noise ratio at the receiver. \( \text{SINR}_{\text{max}} \) is the case of maximum capacity, because in real networks, the data rate is determined by the coding and modulation schemes, so it has a maximum achievable value based on the technology deployed.

The SINR can be calculated as:

\[
\text{SINR} = \frac{P_t}{N_T + I}
\]

where \( P_t \) is the power of the intended transmitter’s signal when the signal reaches the receiver, \( N_T \) is the power of thermal noise and \( I \) is the combined power of signals from any interfering transmitters.

The Friis transmission equation is used to calculate the transmit power \( P_t \):

\[
P_t(d) = P_t \times G_t \times G_r \times \left( \frac{\lambda}{4\pi d} \right)^\eta \times e^{-\alpha d}
\]

where \( P_t \) is the transmit power, \( G_t \) and \( G_r \) are antenna gains of the transmitting and receiving antenna, \( \lambda \) is the wavelength of the signal, \( d \) is the distance between the transmitting and receiving antenna, \( \eta \) is the path loss exponent, and \( \alpha \) is the attenuation factor due to atmospheric absorption.

C. Interference analysis

mmWave signals are generally less prone to mutual interference due to the directional nature of their transmissions resulting from the use of narrow beamwidth antennas. The highly directional signals, together with the blockage effect and the many obstacles in the urban environment, reduce the overall impact of interference. However, secondary mutual interference could still exist when multiple physical links transmit and receive simultaneously in close proximity. For urban wireless backhaul networks, nodes are likely to be relatively high up on the sides or tops of buildings. In this situation, mutual interference has more chance to appear since the wireless nodes have wider LoS view compared with other scenarios such as the street canyon model [28]. Mutual interference must be handled carefully; otherwise, the throughput performance of a multi-hop relaying path could be significantly degraded.

Our relay path construction algorithms consider mutual interference and guarantee that no two physical links along the
constructed path interfere with each other. On the other hand, because relays are simple devices without advanced capabilities such as multiple radio chains, we assume that they are subject to the primary interference constraint meaning that no relay can transmit and receive simultaneously. This simplifies mutual interference avoidance as consecutive physical links are guaranteed not to interfere with each other. This also increases the distance separation between two physical links that could be active at the same time, thereby also lessening interference effects. However, secondary mutual interference is applied on each wireless node (i.e., BS and relay), due to the concurrent transmissions of different wireless links in the network. Mutual interference could exist between any pair of physical links in the network, including BS-to-BS, BS-to-relay, and relay-to-relay physical links.

![Fig. 3. Interference conditions](image)

Fig. 3 shows different interference cases between two disjoint physical links: a) shows the highest interference case, where the interference signal is amplified by a gain of $G_{\text{high}}$ at both ends; in this case, the angle between the useful link direction and the interference signal direction (i.e., $\alpha_1$ and $\alpha_2$ in the figure) is smaller than the half beam width $\frac{\theta}{2}$. b) shows the medium interference case, where the interference signal is only amplified by $G_{\text{high}}$ at one end, and c) shows the lowest interference case, where the interference signal experiences a gain of $G_{\text{low}}$ at both ends; in this case, the interference strength becomes extremely low due to the rapid attenuation of mmWave signals, and we consider this case to be mutual interference-free.

However, in the multi-path construction problem, where a pair of physical links can transmit simultaneously to the same BS, the analysis of mutual interference changes. As shown in Fig. 3d, when two antennas at the sharing node are receiving at the same time, it is the medium interference case, similar to case Fig. 3b. However, when the angle $\alpha$ between the two physical links becomes larger than a threshold value $\beta/2$ (see Fig. 3d), the antenna gain of the side lobe which receives the interference signal becomes extremely small, and this case can also be considered to be interference-free.

1We assume that BSs are not subject to the primary interference constraint imposed on relays, because they are complex enough to support multiple antennas.

IV. PROBLEM FORMULATION

In this section, we formulate the problem of multiple path construction for mmWave backhaul relay-assisted networks using 0-1 integer programming. Specifically, we want to select relay locations from a set of candidate locations to construct multiple paths between a pair of BSs, such that the paths do not interfere with each other and where a relay can be used in only one path.

Given is a network that consists of a set of BSs $B$, a set of logical links $L$ and a set of candidate relay locations $R$. Logical links $L$ constitute the mmWave backhaul topology, which we assume is given. Each logical link $l \in L$ is defined by two elements $(s_l, d_l)$, where $s_l, d_l \in B$ denote the two endpoint BSs of logical link $l$. In this paper, we consider only the problem of selecting relay locations to implement one logical link of the overall mmWave backhaul. Due to the ultra high data rate requirement for backhaul links, we want to deploy multiple parallel paths for a single logical link to achieve the highest possible data rate.

Though the idea of selecting relays might seem simple, selecting physical links to construct the paths is actually more straightforward. Physical links have directions, therefore, it is easier to construct a path from source to destination by selecting physical links. Furthermore, each physical link is unique in our graph model, it is strictly constrained by the starting and ending nodes, however, relays are hard to constrain, they can be used in multiple different physical links. Therefore, we formalize the problem from the perspective of selecting physical links instead of selecting relays.

Using the 3D model of our urban environment and with BSs and candidate relay locations given, we can pre-evaluate the LoS connectivity between each pair of nodes (BS and relays) and obtain a set of possible LoS directional physical links $I$ in the model. Note that if there exists a LoS link between a pair of BSs, this means there is a path between the BSs without the use of relays. This is a special case and we use $I_{\text{BS}}$ to denote the set of physical links (LoS links) that connect two BSs. The set of physical links that include at least one relay node are denoted by $I_a$. Then, we have that $I = I_{\text{BS}} \cup I_a$. For each pair of nodes that have a LoS link between them, a pair of directional physical links $i$ and $-i$ are included in $I$ ($i = 1, 2, 3, ..., \lceil \frac{\alpha}{2} \rceil$). Physical links $i$ and $-i$ share the same end nodes but have different directions. To track the selection status of these directional physical links, we use a set of binary variables $X = \{x_{l,i}\}$, where $l$ and $i$ represent logic link $l$ and physical link $i$, respectively. Therefore, if physical link $i$ is selected by the logical link $l$, we set $x_{l,i} = 1$; otherwise, $x_{l,i} = 0$. The problem to be solved is to determine values of $X$ that satisfy the constraints detailed in the rest of this section.

Single-source single-destination constraint: To select physical links instead of relays in the network, the “starting” and “ending” physical links in a logical link play the equivalent roles as source and destination BSs do in verifying the boundary of the logical link. Therefore, for each logical link
l, we collect all physical links whose starting node is BS $s_l$ into the first hop candidate set ($\mathcal{U}_l$), similarly, for all physical links whose ending node is BS $d_l$, we collect them into the last hop candidate set ($\mathcal{V}_l$). We use binary matrices $U = \{u_{i,l}\}$ and $V = \{v_{i,l}\}$ to record whether physical link $i$ is in set $\mathcal{U}_l$ and/or in set $\mathcal{V}_l$. If $i \in \mathcal{U}_l$, then $u_{i,l} = 1$; otherwise, $u_{i,l} = 0$. Similarly, if $i \in \mathcal{V}_l$, then $v_{i,l} = 1$; otherwise, $v_{i,l} = 0$.

For each logical link $l$, it must contain the “starting” and “ending” physical links, therefore, each logical link has one physical link from $\mathcal{U}_l$ and one physical link from $\mathcal{V}_l$, which can be formulated as follows\(^2\):

$$\sum_{i \in \mathcal{I}} u_{i,l} x_{t,i} = 1, \sum_{i \in \mathcal{I}} v_{i,l} x_{t,i} = 1, \forall l \in \mathcal{L}. \quad (4)$$

If a logical link $l$ contains only one hop, it means there exists a LoS physical link between this pair of BSs and this LoS link is selected. However, as mentioned before, due to the blockage effect, LoS links will often be unavailable between the BSs in obstacle-rich urban environments. Therefore, the logical links often contain several hops instead of one, and the formulation becomes more complicated when logic links have more than one hops.

**Consecutive link constraint:** When constructing paths with more than one hop, the nodes forming consecutive links along the path must satisfy the following constraint. A pair of physical links $i,j \in I_n$ is “consecutive” if $i$’s ending node is the same as the starting node of $j$. We use a binary indicator matrix $C = \{c_{i,j}\}$ to record the “consecutive” relationship between any pair of physical links in $I_n$. If $i$ and $j$ can be consecutive physical links, $c_{i,j} = 1$; otherwise, $c_{i,j} = 0$. Obviously, if $c_{i,j} = 1$, then $c_{j,i} = 1$. To prevent multiple relays at the same location and account for latency considerations, we constrain that each relay node can only be selected once. Therefore, we need to avoid loops when selecting physical links and, thus, we set $c_{i,-i} = 0, \forall i \in I_n$.

It is also obvious that the number of consecutive link pairs contained in a logical link is one less than the number of physical links in the logical link.

Note that for the physical links in the first hop candidate set, there is no physical link preceding them, and for the physical links in the last hop candidate set, there is no physical link following them. This is because the paths that we consider start from the source BS and end at the destination BS. Therefore, there is no physical link “before” the source BS and there is no physical link “after” the destination BS. Hence, we have,

$$u_{l,i} c_{j,i} x_{t,j} x_{t,i} = 0, \forall i, j \in I_n, \forall l \in \mathcal{L}. \quad (5)$$

$$v_{l,i} c_{j,i} x_{t,j} x_{t,i} = 0, \forall i, j \in I_n, \forall l \in \mathcal{L}. \quad (6)$$

Furthermore, to ensure all the selected links construct a well connected path, we need to make sure that: 1) any selected source link that is in the first hop candidate set has a consecutive link after it, 2) any selected destination link that is in the last hop candidate set has a consecutive link before it, and 3) any other physical link has consecutive links before and after it. This constraint can be formulated as:

$$u_{l,i} c_{j,i} x_{t,j} x_{t,i} = 1, \forall i, j \in I_n, \forall l \in \mathcal{L}. \quad (7)$$

$$v_{l,i} c_{j,i} x_{t,j} x_{t,i} = 1, \forall i, j \in I_n, \forall l \in \mathcal{L}. \quad (8)$$

$$c_{i,j,k} x_{t,i} x_{t,j} x_{t,k} = 1, \forall i, j, k \in I_n, \forall l \in \mathcal{L}. \quad (9)$$

**Dedicated relay constraint:** Each relay node can only be selected once, and each physical link (except LoS links) contains at least one relay node. Therefore, we constrain that each physical link can only be selected at most once, which can be formulated as:

$$\sum_{l \in \mathcal{L}} x_{l,i} \leq 1, \forall i \in \mathcal{I}. \quad (10)$$

As logical links are directional, if physical link $i$ is selected by a logical link $l$, to avoid using repeated relay nodes, physical link $-i$ cannot be selected by any logical link in the network. Therefore, we have

$$\sum_{l \in \mathcal{L}} x_{l,i} + x_{l,(-i)} \leq 1, \forall i \in \mathcal{I}. \quad (11)$$

Similarly, to avoid using repeated relay nodes, if two physical links belong to a same consecutive link pair, they cannot be selected by different logical links. Thus,

$$c_{i,j} x_{k,i} x_{j,i} = 0, \forall i, j \in I_n, k \neq l \in \mathcal{L}. \quad (12)$$

**In-and-out-once constraint:** To avoid generating loops when relays are selected to construct a logical link, we restrict that each relay must have exactly one inward physical link and one outward physical link, which is defined as the “in-and-out-once” constraint. However, from the perspective of selecting physical links, it is hard to describe this constraint with relay nodes. Instead, the concept of “consecutive” link pairs can be utilized to describe this constraint. We define that the “front” link must be followed by the “back link” in the consecutive link pair. The following theorem describes a limitation on links designated as front and back links in order to prevent loops in the logical paths.

**Theorem 1.** To prevent the generation of loops when constructing logical link $l$, the physical link $i$ which is selected by logical link $l$ (i.e., $x_{l,i} = 1$) can only be used once as a front link and once as a back link among all consecutive link pairs in $l$.

**Proof.** If a logical link $l$ contains a loop, there must exist a node connects with at least three physical links (i.e., a branch-like structure), which means there exists at least two consecutive link pairs at this node, and these consecutive link pairs share either the same front link or the same back link, which is denoted as physical link $i$. Therefore, we constrain that each selected physical link can only be used once as front link and once as the back link among all consecutive link pairs in $l$.

Theorem 1 allows us to formulate this constraint as:

$$\sum_{j \in I_n} (1 - u_{l,i}) c_{j,i} x_{t,j} x_{t,i} \leq 1, \forall i \in I_n, \forall l \in \mathcal{L}. \quad (13)$$

\(^2\)Additions in this section are decimal, not binary, calculations, i.e. $1+1=2$, $1+1 \neq 1$. 

85
\[
\sum_{j \in I_n} (1 - v_{i,j})c_{i,j}x_{l,i}x_{l,j} \leq 1, \quad \forall \ i \in I_n, \ \forall \ l \in \mathcal{L}. \quad (14)
\]

**Mutual interference constraint:** The mutual interference relationship between each pair of physical links in the network can be pre-computed based on the interference analysis in Section III. We use a binary matrix \( \mathcal{M} = \{m_{i,j}\}, \ i, j \in I \) to record the interference relationship between different pairs of physical links. If physical link \( i \) and physical link \( j \) interfere with each other, we set \( m_{i,j} = 1 \); otherwise, \( m_{i,j} = 0 \).

**Intra-flow interference constraint:** For a logical link \( l \), the intra-flow mutual interference may exist between any pair of physical links \( i \) and \( j \), unless they belong to a same consecutive link pair. As we mentioned before, we assume relays are subject to the primary interference constraint which means no relay can transmit and receive simultaneously. Therefore, for the consecutive physical links belong to the same logical link, they cannot be activated simultaneously, thus no intra-flow mutual interference exists between consecutive link pairs. In order to formulate this constraint, we need to guarantee interference link pairs will never be selected by a same logical link simultaneously, unless they belong to the same consecutive link pair. Thus, this constraint can be formulated as
\[
(1 - c_{i,j})m_{i,j}x_{l,i}x_{l,j} = 0, \quad \forall \ i, j \in I_n, \ \forall \ l \in \mathcal{L}. \quad (15)
\]

**Inter-flow interference constraint:** In addition to considering the intra-flow interference from the “self-traffic”, inter-flow interference from multiple difference logical links also needs to be considered. The concurrent transmissions between different physical links in different logical links may generate the so-called inter-flow mutual interference. To formulate this inter-flow interference constraint, we need to specify the fact that if two physical links \( i \) and \( j \) belong to different logical links, then \( i \) and \( j \) do not interfere, i.e. \( m_{i,j} = 0 \). This leads to the following additional constraint:
\[
m_{i,j}x_{l,i}x_{l,j} = 0, \quad \forall \ i, j \in I, \ \forall \ l, k \in L, \ l \neq k. \quad (16)
\]

Our final problem formulation is to determine values of \( \mathcal{X} \) that satisfy the seven sets of constraints given above.

**V. BOOLEAN SATISFIABILITY**

Boolean satisfiability (Bool-SAT) problems are well-known constraint satisfaction problems that appear in many areas of computer science and engineering. The significant improvement in Bool-SAT solving methods has led to the development of many successful SAT solver programs. Here, we formulate our multi-path construction problem in Bool-SAT terms so we can apply modern SAT solvers to it.

The Bool-SAT problem is expressed in conjunctive normal form (CNF), which is also called product-of-sums form. Each conjunction term or each sum term in the CNF is called a clause, and each clause is a disjunction of literals, where a literal is either a variable or the negation of a variable. In order for the entire formula to evaluate to 1, each clause must be satisfied, i.e. evaluate to 1. For example, the CNF formula
\[
f = a \land (\neg a \lor b \lor c) \land (\neg b \lor c) \quad (17)
\]
consists of 3 variables and 3 clauses. The formula is satisfiable by the assignment of \( \{a = 1, b = 0, c = 1\} \), whereas the assignment \( \{a = 0, b = 1, c = 1\} \) does not satisfy \( f \). Note that a SAT problem with \( n \) variables has \( 2^n \) possible assignments.

Now, we study how to construct a formula which is true if and only if there exist multiple paths in a given area. To construct multiple paths by selecting relays in a relay set, all the constraints, as mentioned earlier, need to be satisfied. To present this problem into the CNF format, we can say that this problem is satisfiable if and only if each constraint is satisfied individually. Then we can encode each of the constraints to CNF format.

1) Single-source single-destination constraint: For a given pair of BSs, the binary matrices \( U \) and \( V \) are given values, thus, the first hop candidate set \( U_l \) and last hop candidate set \( V_l \) are given information. Assuming that there are 3 physical links \( x_{l,a}, x_{l,b} \) and \( x_{l,b} \) in set \( U_l \). Then, the first equation in Eq. 4 can be simplified as
\[
\sum_{i \in I} x_{l,i} = 1, \quad \forall \ l \in \mathcal{L}. \quad (18)
\]

In this equation, the sum of \( x_{l,i} \) equals one, which means for all the physical links in \( U_l \), only one of them is selected. To encode to the CNF format, each physical links \( x_{l,i} \) can be defined as variable and Eq. 18 can be present in CNF format as
\[
\neg x_{l,a} \lor \neg x_{l,b} = 1, \quad \neg x_{l,a} \lor \neg x_{l,c} = 1,
\]
\[
\neg x_{l,b} \lor \neg x_{l,c} = 1, \quad x_{l,a} \lor x_{l,b} \lor x_{l,c} = 1; \quad \forall \ l \in \mathcal{L}. \quad (19)
\]

Each of these equations can be seen as a clause in our CNF formula. Thus, for different link \( l \), with the given set \( U_l \) and \( V_l \), it can follow the steps from Eq. 18 to Eq. 19, to present this constraint in CNF format.

2) Consecutive link constraint: In a given area, the positions of relays and BSs are pre-determined, therefore, the physical links and the consecutive relationship \( c_{i,j} \) are given parameters. Thus, Eq. 5 and Eq. 6 can be simplified as
\[
x_{l,i}x_{l,j} = 0, \quad \forall \ l \in \mathcal{L}. \quad (20)
\]

where \( x_{l,i} \) belongs to set \( U_l \) or set \( V_l \) and physical link \( i, j \) belong to a same consecutive link pair. To present in CNF format, Eq. 20 can be expressed as
\[
\neg x_{l,i} \lor \neg x_{l,j} = 1, \quad \forall \ l \in \mathcal{L}. \quad (21)
\]

Therefore, each of the equations for different links like Eq. 21 can be defined as clause and \( x_{l,i}, x_{l,j} \) can be defined as variables in CNF formula. For the constraint expressed in Eq. 7, Eq. 8 and Eq. 9, they can follow the same method as Eq. 20. However, there is a special case, assuming in Eq. 7, for a source physical link \( i \), if it does not have a following consecutive link, then the equation can be simplified as
\[
\neg x_{l,i} = 1, \quad \forall \ l \in \mathcal{L}. \quad (22)
\]

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on July 22,2021 at 18:41:29 UTC from IEEE Xplore. Restrictions apply.
3) Dedicated relay constraint: To construct multiple paths in a given area, the number of multiple paths is a given value. Assuming we need to construct two paths in this given area, Eq. 11 can be simplified as
\[ x_{l_1,j} + x_{l_1,(i-1)} + x_{l_2,i} + x_{l_2,(i-1)} \leq 1, \forall i \in \mathcal{I}. \]  
where \( l_1 \) and \( l_2 \) means the two paths, separately. Each of the physical link can be seen as a variable in the CNF format, therefore, in Eq. 23, it constraints that at most one variable can be evaluated to 1 and all other variables have value 0. To present in CNF format, Eq. 23 can be expressed as
\[
\neg x_{l_1,i} \lor \neg x_{l_1,(i-1)} = 1, \\
\neg x_{l_2,i} \lor \neg x_{l_2,(i-1)} = 1, \\
\neg x_{l_1,i} \lor \neg x_{l_2,(i-1)} = 1, \\
\neg x_{l_1,(i-1)} \lor \neg x_{l_2,i} = 1, \forall i \in \mathcal{I}.
\]  
(24)

Each of equations in Eq. 24 can be seen as clauses in CNF format. For Eq. 12, it follows the same step with Eq. 5 to simplify to the CNF format. Therefore, with given number of multiple paths and physical links, the dedicated relay constraint can be presented in CNF format.

4) In-and-out-once constraint: For each physical link \( x_{l,i} \), whether it is selected or not is not given, therefore, in Eq. 13 and Eq. 14, the equations are less or equal than one. These equations can rewrite as
\[
(1 - u_{l,i}) \left\{ \neg x_{l,i} + x_{l,i} \sum_{j \in \mathcal{L}} c_{l,j} x_{l,j} \right\} = 1,
\]
\[
(1 - v_{l,i}) \left\{ \neg x_{l,i} + x_{l,i} \sum_{j \in \mathcal{I}} c_{l,j} x_{l,j} \right\} = 1,
\]
\[
\forall i \in \mathcal{I}, \forall l \in \mathcal{L}.
\]  
(25)

where \( \neg x_{l,i} \) represents the physical links which are not selected. As mentioned earlier, \( u_{l,i} \), \( v_{l,i} \), \( c_{l,j} \) and \( c_{l,i} \) are given parameters. Assuming in a logic link \( l \), the physical link \( i \) is not a "starting" link, it shares a consecutive link pairs with 2 different physical links \( p \) and \( q \), \( i \) is the back link in these link pairs. Then the first equation in Eq. 25 can be simplified as
\[
\neg x_{l,i} + x_{l,i} \sum_{j \in \mathcal{L}} x_{l,j} = \neg x_{l,i} + \sum_{j \in \mathcal{L}} x_{l,j} = 1, \forall l \in \mathcal{L}. \]  
(26)

which forces only one of physical link can be selected from the two available physical links. To present in CNF format, Eq. 26 can be rewritten as
\[
\neg x_{l,i} \lor \neg x_{l,i,p} \lor \neg x_{l,i,q} = 1, \\
\neg x_{l,i} \lor x_{l,i,p} \lor x_{l,i,q} = 1.
\]  
(27)

Each of the equation above can be seen as a clause in the CNF formula. Therefore, following these transformation steps, this constraint can be presented in CNF format.

5) Mutual interference: As mentioned earlier, with the given physical links, the mutual interference relationship between each pair of physical links can be pre-computed and the consecutive relationship are given parameters. Therefore, the values of \( m_{l,i} \) and \( c_{l,i} \) for different link pairs are given. Then, Eq. 15 and Eq. 16 can be simplified as
\[
x_{k,i} x_{l,j} = 0, \forall i, j \in I_{k}, k, l \in \mathcal{L}. \]  
(28)

To present in CNF format, Eq. 28 can be expressed as
\[
\neg x_{k,i} \lor \neg x_{l,j} = 1, \forall i, j \in I_{k}, k, l \in \mathcal{L}. \]  
(29)

Therefore, each of the equations for different links like Eq. 29 can be defined as a clause and \( x_{k,i}, x_{l,j} \) can be defined as variables in CNF formula.

Once we have converted our problem into a Boolean satisfiability problem, we can use a SAT solver to check whether there exist multiple paths in a given area. Our simulation results in later sections use a SAT solver which is based on the Davis-Putnam-Logemann-Loveland (DPLL) algorithm [29], [30]. This is a backtracking-based search algorithm that traverses the variable assignments until a satisfying assignment is found (the formula is satisfiable), or all combinations have been exhausted (the formula is unsatisfiable).

It turns out that when the network size becomes large, i.e. the number of BSs and relays increases, the number of variables and the number of clauses grow rapidly and, therefore, the number of possible assignments to test also increases rapidly. We can predict that, for a SAT solver with a large number of assignments, even with a powerful method that expedites the backtrack search algorithm, a substantial computation time will be required to produce a solution.\(^3\) Thus, in the next section, we propose a heuristic algorithm that terminates much faster than the SAT solver in larger networks.

VI. HEURISTIC ALGORITHM FOR MULTIPATH CONSTRUCTION

In this section, we devise a heuristic algorithm to solve the multiple paths construction problem in relay-assisted mmWave backhaul networks. In our earlier work [9] [10], we developed single path construction algorithms that find optimal-throughput paths in mmWave backhaul networks. However, in the upcoming simulation section, we show that multiple (non-optimal) paths easily outperform single optimal-throughput paths. Moreover, our earlier algorithms cannot be applied directly to the multiple path construction problem. This is because, in addition to eliminating intra-flow interference, we also need to consider inter-flow interference between different paths, which is not considered in the earlier work. In addition to the interference consideration, if multiple paths are constructed independently, the same relay location could be selected in multiple paths. However, this situation is not allowed in our problem formulation, as mentioned earlier.

To restate the problem, we need to find multiple node-disjoint paths in a given area with interference and re-use constraints, using all of the candidate relay locations as potential nodes in the paths. Some earlier works [22] proposed finding multiple paths based on a maximum flow (max-flow) method.

\(^3\)Our simulation results in Sec. VII validate this.
Referring to these earlier works, we can find multiple vertex-disjoint paths by reducing the problem to a max-flow problem in an appropriately constructed graph. After multiple paths are found, we can check pairs of paths for inter-flow interference. In order to maximize performance, we will choose the pair of paths with highest combined throughput among all pairs of interference-free paths.

To construct an appropriate graph for the max-flow problem, we use the “node splitting” method mentioned in [22] as shown in Fig. 4. The objective of node splitting is to transform the maximum flow problem with node-disjoint constraint into the standard maximum flow problem with link capacity constraint. An appropriate graph is constructed as follows:

1) Split each relay \( v \) into two virtual nodes \( v_{\text{in}} \) and \( v_{\text{out}} \). Add a link \((v_{\text{in}}, v_{\text{out}})\) with link capacity 1.
2) Replace each other edge \((u, v)\) in the graph with an edge from \( u_{\text{out}} \) to \( v_{\text{in}} \) of capacity 1.
3) Add in a new dedicated source node \( s \) and destination node \( d \). Replace each source edge with \((s, v_{\text{in}})\) and each destination edge with \((v_{\text{out}}, d)\), and set all edge capacities to 1.

\[
\begin{array}{c}
C_1 \quad V \quad C_2 \\
\end{array}
\]

With these procedures, node \( v_{\text{in}} \) receives all the input flows of relay \( v \), node \( v_{\text{out}} \) sends all the output flows of \( v \), and the added link \((v_{\text{in}}, v_{\text{out}})\) will constrain relay \( v \) to be selected at most once. Since all edges in this graph have a capacity of 1, this means that each physical link can only be used once.

Once we have generated a new appropriate graph from the original graph, we can transform the multiple paths construction problem into a max-flow problem and find multiple paths in the new graph. We then check the resulting paths for interference. By checking interference only after constructing the multiple paths, some interference-free paths might be missed using this method. However, our results in the later section show that this method finds multiple interference-free paths at a rate very close to the SAT solver meaning that, in practice, potential solutions are not missed frequently. In the case where multiple pairs of interference-free paths are produced by the max-flow method, we simply choose the pair with highest combined throughput. The pseudocode for our heuristic algorithm, referred to as Algorithm Max-IFMP, is shown in Algorithm 1.

**VII. NUMERICAL RESULTS AND SIMULATIONS**

In this section, we compare Algorithm Max-IFMP against the optimal single path solution and the SAT-solver method. As mentioned earlier, we use an actual 3D topology of a section of downtown Atlanta to drive the simulations. Within this topology, we deployed 42 BSs at select positions and 183 candidate relay locations in a 1200m × 1600m area. We limit the maximum physical link distance to 300 meters, since longer LoS paths supporting Gbps rates rarely exist in a dense urban environment due to signal attenuation and blockages.

Within this setting, we randomly chose BS pairs separated by a distance in the range of \([20, 200],[200, 400], [400, 600], [600, 800], [800, 1000]\). 100 BS pairs were picked for each distance range. The fixed parameter values mentioned in previous equations are shown in Table I. Due to the short LoS links used in the backhaul and the high SINR at the receiver, we ignore the relatively small random attenuation due to shadowing effects in our analysis. However, implementation loss (5dB), noise margin (5dB), and heavy rain attenuation (10dB/km) are considered in the analysis by subtracting an additional link margin \( L_{\text{att}} = 10 \text{dB} + 10 \text{dB/km} \times d \) when calculating the received power.

### Algorithm 1 Max-IFMP: Finding multiple interference-free paths using a modified max-flow algorithm

**Input:** original network, transformed network \( G = (V, E) \) with all flow capacities equal to 1, a source node \( s \), and a sink node \( t \)

**Output:** A pair of interference-free paths in original graph or empty paths if no solution found

1: Create the residual graph \( G_f \)
2: Initialize flow on all edges to 0
3: while path exists from \( s \) to \( t \) in \( G_f \) do
4: \( c_f(p) = \min (c_f(u, v) : (u, v) \in p) \)
5: add this path to transform – path
6: for each edge \((u, v) \in p\) do
7: \( \text{if} \ (u,v) \text{ is forward edge} \) then
8: \( \text{flow}(u,v) = \text{flow}(u,v) + c_f(p) \)
9: \( \text{else} \)
10: \( \text{flow}(u,v) = \text{flow}(u,v) - c_f(p) \)
11: for each path in transform – path do
12: remove all nodes inserted by the node-splitting method
13: add path to orig – path
14: \{ now, find max. throughput pair of interf.-free paths \}
15: \( P_s = P_t, P_h = \epsilon \)
16: for each \( P_s, P_h \in \text{orig – path} \) do
17: \( \text{if} \ P_s \text{ and } P_h \text{ are interference-free and combined throughput is} \)
18: \( \text{return} \ P_s, P_h \)

### Table I
 **PARAMETERS OF SIMULATION ENVIRONMENT**

<table>
<thead>
<tr>
<th>( B )</th>
<th>2.16 GHz</th>
<th>( P_f )</th>
<th>1 W</th>
<th>( G_s, G_p )</th>
<th>21.87 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
<td>5 mm</td>
<td>( \eta )</td>
<td>2.0</td>
<td>( \text{SINR}_{\text{max}} )</td>
<td>50 dB</td>
</tr>
<tr>
<td>( L_{\text{att}} )</td>
<td>10 dB</td>
<td>( \alpha )</td>
<td>16 dB/km</td>
<td>\beta</td>
<td>30°</td>
</tr>
</tbody>
</table>

### A. Comparison of Single Optimal Path vs. Multiple Paths

In this subsection, we compare the throughput performance of a single optimal relay path against the multiple relay paths produced by Algorithm Max-IFMP. To find the optimal single path, we use our algorithm from [10]. Fig. 5 shows the average throughputs that are produced for these two different cases. The figure clearly shows that finding multiple (non-optimal) paths significantly outperforms a single optimal path. For the
shorter BS separations, which is the common case for densely-deployed SBSs, the throughput is increased by 60–100% when using multiple paths.

We note that the average throughputs for Algorithm Max-IFMP reported in Fig. 5 are calculated only across the cases where a multi-path solution was found. However, there are some cases where the algorithm did not find a solution. To evaluate how frequently this situation occurs and to determine how often Algorithm Max-IFMP fails to find a solution when multiple interference-free paths do exist, we compare it against the SAT-solver method in the next subsection.

B. SAT-Solver vs. Heuristic algorithm

1) Feasibility of Multiple Paths: Here, we evaluate the feasibility of finding multiple paths for different BS separations in different sized areas with the SAT-solver method and Algorithm Max-IFMP. We considered two BS separations, [20, 200) and [200, 400), and varied the size of the area considered. Within our total area of 1200m × 1600m, we selected three different subareas of sizes 400m × 400m, 600m × 600m and 800m × 800m, respectively. Fig. 6 shows the results for these cases. The satisfaction rate is defined as the fraction of cases for which multiple paths between BSs that satisfy all the relevant constraints were found. Note that the satisfaction rate increases as the size of the area increases. With an area of size 400m × 400m and a BS separation in the range [200, 400), the SAT solver is able to find multiple paths about 80% of the time, whereas with an area of size 600m × 600m and the same BS separation, it finds multiple paths in every case. Note also that Algorithm Max-IFMP produces a satisfaction rate that is very close to that of the SAT-solver method for all area sizes and BS separations in Fig. 6. Thus, while Algorithm Max-IFMP does not find multiple paths in all cases where it is possible, in practice it is very close to the perfect solution (in terms of feasibility) given by the SAT solver.

2) Running Time Comparison: As mentioned earlier, the running time of the SAT solver is likely to increase rapidly as the size of the area grows. This is because a larger area will provide more candidate relay locations, and thus more possible physical links to consider, since the number of possible physical links grows quadratically with the number of candidate relay locations. This will rapidly increase the numbers of variables and clauses, which will greatly impact the running time of the SAT-solver. Fig. 7 shows the running times for BS separations in the ranges of [20, 200) and [200, 400) meters and areas of size 400m × 400m, 600m × 600m, and 800m × 800m. Note that, as expected, the running time for the SAT-solver does increase very rapidly as the area size increases. For example, while both the SAT-solver method and Algorithm Max-IFMP have 100% satisfaction rate in the separation range of [20, 200) meters with 600m × 600m area, the SAT-solver required more than 2 minutes to find multiple paths for each pair of BSs, whereas Algorithm Max-IFMP only needed 0.255 seconds to obtain the paths. For larger areas, e.g. 800m × 800m or higher, the running time needed for the SAT-solver is a few hours to tens of hours, whereas Algorithm Max-IFMP handles even very large sizes in a few seconds. Given the small difference in satisfaction ratio, this performance of Algorithm Max-IFMP is quite good.

3) Throughput Comparison: The SAT-solver only finds a pair of interference-free paths when possible, but it does not factor in the throughputs of the paths in its formulation. However, when more than one pair of the paths found by the max-flow algorithm are found to be interference free, Algorithm Max-IFMP chooses the pair with the highest combined throughput. Thus, in many cases, the aggregate throughput of the paths found by Max-IFMP exceeds that of the SAT-solver solution. Fig. 8 compares the average aggregate throughputs of the solutions found by the two approaches. The average throughput increase for Max-IFMP is 5–10%.

Combining all the results in this subsection, we see that,
for a small sacrifice in satisfaction ratio. Algorithm Max-IFMP finds higher-throughput paths than the SAT solver while taking only a small fraction of the running time.

The SAT-solver method and the heuristic Algorithm Max-IFMP each have their own merits. The SAT solver can give a precise judgment on whether multiple paths exist, but it can take a very long time to produce a result, especially when the number of candidate relay locations is large. Algorithm Max-IFMP does not guarantee to find multiple interference-free paths when such paths exist, but it finds multiple paths efficiently and factors aggregate throughput into its path selection. In practice, these two methods could work together: first, Algorithm Max-IFMP can be executed and second, only when Max-IFMP fails to find a solution, the SAT-solver method can be used to check if a solution exists. In this way, the long execution time penalty of the SAT solver is only incurred when Max-IFMP fails and we can guarantee to find multiple interference-free paths whenever they exist.

VIII. CONCLUSION

In this paper, we have presented algorithms for multi-path construction in relay-assisted mmWave backhaul networks. Through simulations of a 3-D topology from downtown Atlanta, we evaluated the feasibility and benefits of multi-path construction. The results showed that a fast heuristic multi-path construction algorithm substantially outperforms an optimal-throughput single path algorithm. While the heuristic is not always guaranteed to find a multi-path solution when one exists, comparisons against a SAT solver show that it does construct multiple interference-free paths in the vast majority of cases where it is possible to do so. Although not explored herein, having multiple paths also provides robustness to blockages and failures along one of the paths. A detailed evaluation of this benefit is left for future work.

ACKNOWLEDGEMENT

This research was supported in part by the National Science Foundation through Award CNS–1813242.

REFERENCES
