Autonomous Relay for Millimeter-Wave Wireless Communications
Linghe Kong, Member, IEEE, Linsheng Ye, Fan Wu, Member, IEEE, Meixia Tao, Senior Member, IEEE, Guihai Chen, and Athanasios V. Vasilakos, Senior Member, IEEE

Abstract—Millimeter-wave (mmWave) communication is the rising technology for next-generation wireless transmission. Benefited by its abundant bandwidth and short wavelength, mmWave is advanced in multi-gigabit transmittability and beamforming. In contrast, the short wavelength also makes mmWave easily blocked by obstacles. In order to bypass these obstacles, relays are widely needed in mmWave communications. Unmanned autonomous vehicles (UAVs), such as drones and self-driving robots, enable the mobile relays in real applications. Nevertheless, it is challenging for a UAV to find its optimal relay location automatically. On the one hand, it is difficult to find the location accurately due to the complex and dynamic wireless environment; on the other hand, most applications require the relay to forward data immediately, so the autonomous process should be fast. To tackle this challenge, we propose a novel method AutoRelay specialized for mmWave communications. In AutoRelay, the UAV samples the link qualities of mmWave beams while moving. Based on the real-time sampling, the UAV gradually adjusts its path to approach the optimal location by leveraging compressive sensing theory to estimate the link qualities in candidate space, which increases the accuracy and save the time. Performance results demonstrate that AutoRelay outperforms existing methods in achieving an accurate and efficient relay strategy.

Index Terms—Millimeter-wave communications, mobile relay, compressive sensing.

I. INTRODUCTION

The next-generation wireless technology, millimeter-wave (mmWave) communication [19], show its great potential on emerging networks, such as 5G cellular networks [27], wireless personal area network WiGig [20], Internet of Things [11], and wireless data center [8]. The promising mmWave communication operates at 3-300GHz and its available bandwidth is up to multiple GHz. Thus, mmWave is able to provide multi-gigabit transmittability [7] for wireless information shower [19]. Moreover, mmWave’s short wavelength allows multiple antennas to be integrated in a tiny space. Consequently, beamforming technique [20] can be exploited to construct directional beam, prolong the communication range, and reduce the interference.

Although mmWave possesses the above promises, it has two drawbacks. First, mmWave’s communication range is shorter than existing wireless technologies such as WiFi and Bluetooth, because the signal with shorter wavelength attenuates more severely in air [6]. Second, mmWave is easily blocked by obstruction. Since the transmission power is limited to 40dBm by Federal Communications Commission, mmWave has inadequate power to penetrate through most obstacles [27].

To compensate these disadvantages, relays [13], [15], [21] are essential to prolong the communication range and bypass the obstacles. With the rapid development of unmanned autonomous vehicles (UAVs) [10], [24], mobile relays such as drones and self-driving have become possible in real applications, especially in emergent and critical tasks.

Mobile relays have raised great interests in both industry and academia due to its low-cost and flexibility. In literature, there are lots of related studies for relays in conventional wireless techniques such as WiFi and cellular [4], [22]. In addition, cooperative UAV relays are investigated [14], [25] to further increase the connectivity and throughput in mobile networks. However, as an emerging wireless technique, mmWave communication has nearly no tailored study on mobile relays.

It is challenging for a UAV to find its optimal location accurately and quickly. Since the wireless environment is dynamic and unpredictable, a UAV has no knowledge to determine the optimal location easily. An accurate method is to measure the link qualities between transmitter and receiver in all possible space and select the optimal one. Nevertheless, this method costs too long time while most applications require the relay to start working as soon as possible.

To address this problem, we propose a novel solution Autonomous Relay (AutoRelay). Different from determining the relay location by theoretical models, AutoRelay measures the real-time link qualities while moving. The core design of AutoRelay takes advantage of compressive sensing to measure only a few samples and online estimate the link qualities in whole candidate space. Hence, the UAV can automatically adjust its path to approach the optimal relay location quickly. AutoRelay has two advantages: i) Combining the mmWave model, real-time sampling, and compressive sensing theory, AutoRelay can estimate the optimal location accurately. ii) Leveraging the online update and continual approaching mechanism, AutoRelay can quickly arrive the relay location.

We further develop AutoRelay for more realistic applications. The enhanced AutoRelay can not only construct directional beam, prolong the communication range, and reduce the interference.

Although mmWave possesses the above promises, it has two drawbacks. First, mmWave’s communication range is shorter than existing wireless technologies such as WiFi and Bluetooth, because the signal with shorter wavelength attenuates more severely in air [6]. Second, mmWave is easily blocked by obstruction. Since the transmission power is limited to 40dBm by Federal Communications Commission, mmWave has inadequate power to penetrate through most obstacles [27].
serve the simple one-transmitter one-receiver one-relay scenario but also extend to complex scenarios including multiple transceivers and multiple relays. In addition, mmWave usually adopts directional beam to increase the communication range. The enhanced AutoRelay could operate well in both omnidirectional and directional communication systems.

In order to evaluate the performance of AutoRelay, we conduct extensive simulations. In our simulations, we assume the mobile relay is a drone, which could move freely in 3D space. Performance results demonstrate that AutoRelay outperforms the classic and the state-of-the-art relay methods in terms of accuracy and efficiency.

The contributions of this paper are summarized as follows:

- To the best of our knowledge, this is the first work focusing on autonomous mobile relays for millimeter-wave communications.
- A novel AutoRelay is designed to address the relay problem in mmWave communications. Benefited from the compressive sensing theory, AutoRelay drives the UAV to the optimal relay location accurately and quickly.
- Extensive simulations are conducted to evaluate AutoRelay in different cases. Performance results demonstrate the advantages of AutoRelay.

The rest of the paper is organized as follows. Section II introduces the background. Section III formulates the system models and states the problem. In Section IV, we design the basic AutoRelay. In Section V, we improve AutoRelay for more realistic applications. In Section VI, we conduct simulations to evaluate AutoRelay. Section VII discusses the related work. Our work is concluded in Section VIII.

### II. BACKGROUND

The study of autonomous relays for millimeter-wave (mmWave) communications is at exactly the right time. On one hand, the mmWave communications do need relay due to its characteristics. On the other hand, recent UAV technique is qualified to realize mobile relays. In this section, we introduce the backgrounds of mmWave communications and UAVs.

#### A. Millimeter-Wave Communications

The next-generation mobile technology, mmWave [19], is envisioned to offer multi-gigabit wireless service for emerging applications. Before designing mobile relays for mmWave, we first introduce the advantage and disadvantages of mmWave.

The first promise of mmWave is the bandwidth. Take 60GHz as an example. The unlicensed 60GHz band provides 7GHz bandwidth for mobile applications and is supported by IEEE 802.11ad standard, targeting indoor multi-gigabit wireless networks. Benefiting from the wide bandwidth, the bitrate of 802.11ad is up to 6.76Gbps [7]. The other key parameters in 802.11ad are listed in Table I.

Besides the bandwidth, another promise is the short wavelength. Since the wavelength of mmWave is at millimeter level, it is possible to pack a large number of antennas into a small space. e.g, a 100-element 60GHz array can be integrated into one square inch. Thus, the beamforming technique can be applied in mmWave. Beamforming [20] is a signal processing technique to generate directional signal transmission by smart antenna array. Although the transmission range of mmWave is only 20m in omni-directional broadcast mode, beamforming can concentrate power into one direction and offer a transmission range that exceeds 130m.

However, the characteristics of mmWave is significantly different from conventional wireless techniques such as WiFi, Bluetooth, WiMax, and 4G (see Table I). The disadvantages of mmWave include its short propagation and low penetration.

- Propagation. In free space, the loss of 60GHz mmWave is about 16dB/km [6]. Although mmWave is hard to establish a kilometer-level link, it works well within a short range because beamforming is able to enhance the spatial reuse. For example, 100m using beamforming.
- Penetration. While 2.4GHz signals penetrate through some solid materials, mmWave signals are easily blocked by most objects. For example, a human body results in 20-50dB of loss on mmWave. Moreover, since the transmission power is limited to 40dBm by Federal Communications Commission, mmWave does not have adequate power to penetrate obstacles [27].

To avoid the shortcomings, it is envisioned that relays would be largely used in mmWave communications, which could effectively prolong the coverage and bypass the obstacles.

#### B. Unmanned Autonomous Vehicles

Plenty of unmanned autonomous vehicles (UAVs) [10], [24] have been exploited in real applications.

For instance, about 50 Google driverless cars are self-driving in California and Texas for field test. Drones [24], [28] are used for urban sensing and mobile radar imaging. The autonomous robotic fish for ocean debris monitoring [23] can move in the sea and have wireless module to communicate with other robotic fishes or neighboring ships.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>mmWave</th>
<th>WiFi</th>
<th>Bluetooth</th>
<th>WiMax</th>
<th>4G</th>
</tr>
</thead>
<tbody>
<tr>
<td>57-64GHz</td>
<td>2.4/5.8GHz</td>
<td>2.4GHz</td>
<td>2-6GHz</td>
<td>1880-2650MHz</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>802.11ad</td>
<td>802.11a/b/g/n</td>
<td>802.15.1</td>
<td>802.16e</td>
<td>LTE</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>2.16GHz</td>
<td>20</td>
<td>40MHz</td>
<td>1MHz</td>
<td>1.75-20MHz</td>
</tr>
<tr>
<td>Bitrate</td>
<td>693Mbps-6.76Gbps</td>
<td>6-600Mbps</td>
<td>1-24Mbps</td>
<td>Peak upload: 56Mbps, Peak download: 128Mbps</td>
<td>Upstream: 75Mbps, Downstream: 300Mbps</td>
</tr>
<tr>
<td>Tx Range</td>
<td>Directional: &lt;130m, Omnidirectional: &lt;20m</td>
<td>&lt;100m</td>
<td>&lt;100m</td>
<td>&lt;10km</td>
<td>&lt;2km</td>
</tr>
</tbody>
</table>
The Waalbot [17] robot is capable to climb on the wall and ceil. Furthermore, lots of drone platforms allow developers to design their own moving strategies.

Hereby, using UAVs equipped with mmWave communication systems to realize mobile relays is practical recently.

III. PROBLEM STATEMENT

Due to short propagation and low penetration, mmWave transmitter and receiver usually suffer the disconnection problem. Therefore, the mobile relay is needed to tame this problem. In this section, we formulate the mmWave system models, describe the mobile relay problem, and finally take an overview of the compressive sensing on 3D matrix.

A. System Model

To make it easy to understand, our basic system includes only one transmitter (Tx) and one receiver (Rx). As shown in Fig. 1, this Tx transmits mmWave signal to Rx by an omnidirectional manner. Due to the short wavelength, it is common that the mmWave transmission is blocked by a certain obstacle. Hence, a relay is needed to connect Tx and Rx. With the mobility ability, a mobile relay can find an appropriate location to forward the transmission. Note that the system with multiple transmitters, multiple receivers, multiple relays, and directional beam will be discussed in Section V.

1) Space Model: Taking the mmWave transmitter as the origin \( o \) with coordinates \((0, 0, 0)\) and the mmWave receiver as the point \( c \) with coordinates \((n_1, n_2, n_3)\), we can set up a 3D space from \( o \) to \( c \). The optimal relay location is somewhere in this 3D space.

Accordingly, a 3D matrix \( Q \in \mathbb{R}^{n_1 \times n_2 \times n_3} \) is created to represent the link qualities in which \( n_1, n_2, n_3 \) are the scales of the three dimensions. In this way, we split the space into \( n_1 \times n_2 \times n_3 \) small cubes and assume the link quality in one certain cube is the same. The value of each cube \( Q(i, j, k) \) of the matrix denotes the link quality of mmWave transmission if the relay is set in \((i, j, k)\).

2) Link Quality Model: In order to measure the link quality for each location, we provide a model for mmWave transmission, taking the large-scale path loss, the shadowing loss and the non-shadowing loss into consideration [1], [19]. Let \( d \) be the distance and \( P(d) \) be the large-scale path loss between the transmitter and the receiver. \( Z^{(1)} \) and \( Z^{(2)} \) represent the shadowing loss and the non-shadowing loss, respectively.

The loss of the signal \( L \) along the transmission paths can be calculated as
\[
L = P(d) + Z^{(1)} + Z^{(2)},
\]
where \( P(d) \) is given by
\[
P(d) = 10\alpha \log(d) + \beta,
\]
where \( \alpha \) and \( \beta \) are constants and \( \alpha \geq 2 \). Furthermore, considering that attenuation parameter is not constant in the space, using the linear integral over the transmission path, for the shadowing loss \( Z^{(1)} \), we have
\[
Z^{(1)} = \int_{\text{path}} g(r)dr,
\]
where \( r \in \mathbb{R}^3 \) denotes a distance vector in the space and \( g(r) \) dB/m denotes the attenuation parameter due to the shadowing loss on location \( r \). And for the non-shadowing loss \( Z^{(2)} \), we assume a wide-sense stationary Gaussian process with zero mean and variance \( \eta^2 \).

To find a way of link quality evaluation, we should consider both the source side and the destination side. The obvious method to describe the link quality is the received transmission quality from certain points. Since the transmission quality at the destination point received from the relay point is equivalent to that at the relay point received from the destination point, we consider the receiving qualities at the relay point from both the source point and the destination point as factors in representation of link quality metrics. Let \( S_o, S_c \) denote the sending quality at the origin \( o \) and the receiver \( c \) and \( R_o(i, j, k), R_c(i, j, k) \) denote the receiving quality from point \( o \) and \( c \) at point \((i, j, k)\). According to the data transmission model above, we give the representation of the link quality \( Q(i, j, k) \) described as the product of the receiving qualities \( R_o(i, j, k) \) at point \((i, j, k)\) from origin point \( o \) and \( R_c(i, j, k) \) at point \((i, j, k)\) from receiver \( c \)
\[
Q(i, j, k) = R_o(i, j, k) \times R_c(i, j, k),
\]
where \( R_o(i, j, k) \) and \( R_c(i, j, k) \) are given by
\[
R_o(i, j, k) = S_o - L_o,
\]
\[
R_c(i, j, k) = S_c - L_c,
\]
where \( S_o, S_c \) are the sending quality at \( o \) and \( s \) and \( L_o, L_c \) are the signal loss along the path from point \( o \) and \( c \), respectively.

To achieve the optimal transmission quality for the relay, we adopt the product of these two values as the metric.

Let \( H \) denote the set of locations covered by the transmission path and the shadowing loss of the signal is given by
\[
Z^{(1)} = \sum_{(x, y, z) \in H} g(x, y, z)r(x, y, z),
\]
where \( g(x, y, z) \) is the corresponding attenuation parameter of the cube \((x, y, z)\) and \( r(x, y, z) \) is the path length in this cube.

3) Mobile Relay System: Recent UAV technique is adequate to build the mobile relay system. Fig. 2 exhibits an example of mobile relays. Each relay includes an iRobot, a laptop, and a USRP. The iRobot plays the role of a UAV, which could move on a 2D surface. The laptop could control the moving strategy of iRobot. USRP is a software defined radio.
platform to provide the wireless communications. There are also some other mobile relay examples, such as autonomous robotic fish [23] and mmWave drone [28].

B. Problem Description

The mobile relay aims to find the optimal location with the best link quality. We need to develop a strategy of path choosing for the relay and a method to update the quality matrix $Q$ in an online manner.

In this problem, the mobile relay flies along the path according to the strategy in the space with the initialization of the quality matrix $Q$, using the data transmission model. Thus, the problem is described as follows:

The space model is based on the transmitter $o(0,0,0)$ and the receiver $c(n_1, n_2, n_3)$. Given the start point of mobile relay $s(x_s, y_s, z_s)$ and the initial quality matrix $Q$, design a strategy for the relay to find the optimal point $p(x_s, y_s, z_s)$ so that the link quality $Q(x_s, y_s, z_s)$ is maximized.

The link quality calculated by the theoretical model exists errors in real world because noise, multi-path, dynamics, and other factors cannot be perfectly modeled. Hence, the online measurement and update is necessary to correct the theoretical results. However, measuring the link quality at all locations of candidate space is a huge project, which costs too much time. So the proposed AutoRelay only samples at a few locations and estimates other locations using compressive sensing.

For the reason that such strategy is applied in the emergency situations such as fire-fighting, the energy consumption for the relay is not considered. In addition, due to the development of UAV technology, the stability of the mobile relay is not considered as well.

C. Compressive Sensing for 3D Matrix

Compressive sensing [5] is an advanced theory to recover sparse signals by just a few samples. One branch of compressive sensing is matrix completion [3], which has been widely adopted in various applications [12], [16].

To solve our problem, we first introduce some concepts in the matrix completion. For a 3D matrix $X$, $X^{(i)}$ denotes the $i$th frontal slice of $X$. We use $X(:,:,k)$, $X(:,j,:)$ and $X(i,:,,:)$ to denote the $k$th frontal slice, $j$th lateral slice and $i$th horizontal slice, respectively. $\hat{X}$ denotes the 3D matrix obtained by taking the Fourier transform on $X$ along the third dimension.

1) Linear Operation: T-product is used to define the multiplication for 3D matrix. For a 3D matrix $X \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we can view it as an $n_1 \times n_2$ matrix with each element as a tube. Each tube of the 3D matrix can be described as $x(i,j,:)$. Then, we give the definition of t-product.

**Definition 1 (T-Product):** The t-product $C$ of $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $B \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is a matrix of size $n_1 \times n_4 \times n_3$ where the $(i,j)_{th}$ tube denoted by $C(i,j,:) = \mathbf{b}(k,j,:)$ for $i = 1, 2, ..., n_1$ and $j = 1, 2, ..., n_4$ of the matrix $C$ is given by $\sum_{k=1}^{n_2} A(i,k,:)* B(k,j,:)$. The t-product of $A$ and $B$ can be calculated by performing the fast Fourier transformation (FFT) along the tubes of $A$ and $B$ to obtain $\hat{A}$ and $\hat{B}$, then multiplying each pair of the frontal slices of $\hat{A}$ and $\hat{B}$ to obtain $\hat{C}$, and finally performing the inverse FFT along the third dimension to obtain the result. The details are shown in [9].

2) T-SVD: Using multiplication for 3D matrix, we can compute a tensor-Singular Value Decomposition (t-SVD). First, we give the following definitions.

**Definition 2 (Matrix Transpose):** For a matrix $X$ with size $n_1 \times n_2 \times n_3$, $X^\top$ is obtained by transposing each of the frontal slices and reversing the order of transposed frontal slices 2 through $n_3$.

**Definition 3 (Identity Matrix):** The identity matrix $I$ in $\mathbb{R}^{n \times n \times n}$ is a matrix whose first frontal slice is the $n \times n$ identity matrix and all other frontal slices are zero.

**Definition 4 (Matrix Tubal-Rank):** The matrix tubal-rank of a 3D matrix is the number of non-zero tubes of the matrix.

Using t-SVD, we can extract notions of data complexity in the matrix in terms of “rank”. The notion of multi-rank was proposed in [9] using the Fourier Domain representation of t-SVD as the vector of ranks of the slices $\hat{X}(:,:,k), k = 1, 2, ..., n_3$. The $l_1$ norm of the multi-rank can then be a way to measure the complexity of the data in the matrix.

3) Matrix Completion: is used to predict the missing data of a matrix with some of its data being sampled. Suppose there is an unknown matrix $M \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ having a low tubal-rank, which indicates that the elements have high relativity in continuous space. A subset of entries $\{M_{i,j,k} : (i,j,k) \in \Omega\}$ is sampled, where $\Omega$ is an indicator matrix of size $n_1 \times n_2 \times n_3$ meaning which entries of the matrix are being observed. The goal is to recover the entire matrix $M$ from sampled data. To address the problem of matrix completion, the following minimization problem should be solved

\[
\min ||X||_{\text{MNN}}, \text{ subject to } P_{\Omega}(X) = P_{\Omega}(M),
\]

where $||X||_{\text{MNN}}$ denotes the matrix-nuclear-norm (MNN) and is defined as the sum of the singular values of all the frontal slices of $\hat{X}$ [26] and $P_\Omega$ is the orthogonal projector indicating the sampled location. Therefore, the component $(i,j,k)$ of $P_{\Omega}(X)$ equals to the component $(i,j,k)$ of $M$ if $(i,j,k) \in \Omega$ and zero otherwise.
IV. CORE DESIGN OF AUTO RELAY

Based on the compressive sensing, we propose a novel method, Auto Relay, to solve the problem. In this section, we first take an overview of Auto Relay and then show its details.

A. Overview

1) Initial Setup: We assume that points $o$, $c$, $s$ are known initially. These points are obtained by the mobile using the camera to observe the location of the transmitter and the receiver. Then we set up the corresponding 3D space taking the transmitter as the origin $o(0,0,0)$ and the location of the receiver as $c(n_1,n_2,n_3)$. Accordingly, the initial point $s(s_x,s_y,s_z)$ where the mobile relay starts is also set up.

Second, the initial values for the 3D quality matrix $Q$ are set. Considering that the shadowing loss and the non-shadowing loss of the signal are relatively less compared to the large-scale path loss between the sender end and the receiver end, we initialize the quality matrix with values only taking the large-scale path loss into account leading to $L = P(d)$. Note that any form of the distortion on the transmission link quality resulted from the dynamic environment is corrected in the matrix update using the sample real data, so that the initialization does not consider the distortion or the attenuation. Thus, by (4) and (5), for each element in the quality matrix, the value is initialized as

$$Q(i,j,k) = R_o \times R_c = (S_o - L_o) \times (S_c - L_c) = (S_o - P(d_o)) \times (S_c - P(d_c)), \quad (9)$$

where $d_o$ and $d_c$ represent the distances from the mobile relay to the transmitter and the receiver, respectively. And the large-scale path loss $P(d_o), P(d_c)$ are given by (2). The distances $d_o$ and $d_c$ are given by $d_o = \sqrt{i^2+j^2+k^2}, d_c = \sqrt{(i-n_1)^2+(j-n_2)^2+(k-n_3)^2}$.

2) Optimal Location Finding Procedure: In this procedure, samples are taken for updating the link quality matrix. To acquire the link quality of each point, signals are sent by both the transmitter and the receiver to the relay. The data contained in the signals includes the information about the sender, such as the original signal strength $S$ at the sender end and the transmission start time. By measuring the signal strength $R$ at the receiver end, we can get the signal loss $L = S - R$ along the path. After getting the receiving quality $R_o$ from the transmitter point $o$ and $R_c$ from the receiver point $c$, the link quality value for the current location is calculated by (5) and the relay moves to the next location.

We develop a strategy for the mobile relay to take samples for updating the quality matrix and finally find the optimal location. Let set $D$ be the data sample set and $d_o$ denote the data sampled at the location $u(u_x,u_y,u_z)$. By the initialization of the 3D matrix, regardless of the shadowing loss and the non-shadowing loss whose influences will be corrected by matrix update using real sample, we have the general estimation of the link quality matrix. This estimation is good enough for an initialization. Thus, the relay follows the below steps:

- Step 1: Initialize the link quality matrix $Q$ based on the theoretical model of mmWave transmission.
- Step 2: Select the optimal location with the metric $\max(R_o,R_c)$.
- Step 3: Fly towards the optimal location while sampling.
- Step 4: Update the link quality matrix $Q$, using the real-time sampling dataset $D$ and compressive sensing theory.
- Step 5: If the current location is the optimal one, the finding procedure is ended. If not, repeat the procedure from Step 2.

After every round of updating the link quality matrix, if the relay doesn’t arrive at the optimal location, it continues to find the optimal location and takes samples along the way. The exact data of each location is used to update the link quality matrix based on the value achieved before. Specifically, the values taken by the relay and the values previously achieved together are the input of the matrix update procedure to make corrections on the achieved quality matrix. In this way, gradually, we achieve a more accurate link quality matrix for optimal location finding.

B. Algorithm

To solve the problem, we design an algorithm named Auto Relay for the mobile relay to find the optimal location. The detail of the algorithm is shown in Algorithm 1. In general, when the current location is not the optimal location, Auto Relay samples the data along the path to the optimal location and updates the link quality matrix using Matrix-Update method which is introduced later. Then Auto Relay finds the current optimal location and determines whether to terminate or to continue the next iteration.

Algorithm 1 Auto Relay

Input: $o, c, s, Q$.  
Output: $p$.

1: $D \leftarrow \emptyset$
2: Main Procedure 
3: while $s \neq p$ do 
4: for every point $u$ along the path from $s$ to $p$ do 
5: $D \leftarrow D \cup d_u$
6: end for 
7: $Q \leftarrow \text{MatrixUpdate}(Q,D)$
8: $s \leftarrow p$
9: $p \leftarrow \arg\max(Q(u_x,u_y,u_z))$
10: end while

Initially, the input points, the starting point $s$, the origin point $o$ and receiver point $c$ are obtained by the camera carried by the relay itself. The 3D space is set up according to $o$ and $c$. As for the initialization of the quality matrix $Q$, the area where the transmission path is shadowed by the obstructions like buildings is considered to have relatively high signal loss so that the locations in such area cannot be the optimal relay location and are not considered in our problem. Note that the set $D$ which records the data that has been sampled till the current time is initialized to be empty as presented in step 1 in the algorithm.
In the main procedure, AutoRelay takes samples along the path and updates the link quality matrix $Q$. As long as that the current location is not the optimal location, the relay continues to fly to the optimal location and samples data for the next round of matrix update. Such procedure is implemented by a while loop from step 3 to step 10 in the algorithm. Note that after each process of flying to the current optimal location, the number of data observed increases.

By taking the data sampled by the relay and the previously achieved values as inputs, we define matrix update procedure $\text{MatrixUpdate}(Q,D)$ in step 7 in the algorithm as a matrix recovery problem [12], [16], which actually makes corrections on the previously achieved value by using the observed value. Considering the matrix recovery problem, for an unknown 3D link quality matrix $Q$, the goal is to recover the whole matrix $Q$ from the selected entries by solving the problem

$$\min \|X\|_{MNN}, \text{subject to } P_{\Omega_1}(X) = P_{\Omega_1}(Q), \quad (10)$$

which is in the same form as (8). Here, all the entries in the matrix are selected and the exact sampled data is used for entries at which locations have been sampled and the previously achieved value is used for other entries.

Let $\gamma$ be the sampled data, we have $\gamma = P_{\Omega_1}(Q)$. Define $\mathcal{F}_3$ and $\mathcal{F}_3^{-1}$ to be the Fourier and inverse Fourier transform along the third dimension and $\mathcal{G} = \mathcal{F}_3 P_{\Omega_1} \mathcal{F}_3^{-1}$ to be an operator. Thus, we have $\hat{\gamma} = \mathcal{G}(\hat{\gamma}')$ where $\hat{\gamma}'$ and $\hat{\gamma}$ are the Fourier transforms of $\gamma$ and $\mathcal{M}$ along the third dimension. Under such construction, the problem (10) can be represented:

$$\min \|\text{blkdiag}(\hat{X})\|_*, \text{subject to } \hat{\gamma}' = \mathcal{G}(\hat{X}), \quad (11)$$

where blkdiag$(\hat{X})$ is a block diagonal matrix whose diagonal blocks are given by $\hat{X}(i)$. Note that $\|X\|_{MNN} = \|\text{blkdiag}(\hat{X})\|_*$. The optimization problem can be rewritten:

$$\min \|\text{blkdiag}(\hat{Z})\|_*, \text{subject to } \hat{\gamma} = \mathcal{G}(\hat{X}), \quad (12)$$

where $\mathcal{L}$ denotes the indicator function. By applying the framework of Alternating Direction Method of Multipliers (ADMM) [2], we can have the following recursion

$$\hat{X}^{k+1} = \arg\min_{\hat{X}} \{\mathcal{L}_{\gamma = \mathcal{G}(\hat{X})} + \langle \hat{\gamma}, \hat{X} \rangle + \frac{1}{2}\|\hat{X} - \hat{Z}^k\|^2 \},$$

$$\hat{Z}^{k+1} = \arg\min_{\hat{Z}} \{\|\hat{X} - (\hat{Z}^k - \hat{\gamma}^k)\|_F \},$$

$$\hat{\gamma}^{k+1} = \hat{\gamma} + (\hat{X}^{k+1} - \hat{Z}^{k+1}).$$

The solution to (14) is given by the singular value threshold [3]. By examining the format of (14), it can be split into $n_3$ minimization sub-problems. Let $\hat{Z}^{k+1,i}, \hat{X}^{k+1,i}, \hat{\gamma}^{k,i}$ denote the $i$th frontal slice of $\hat{Z}^{k+1},\hat{X}^{k+1}$ and $\hat{\gamma}^{k}$. Then (14) can be split into:

$$\hat{Z}^{k+1,i} = \arg\min_{\hat{Z}} \{\frac{1}{2}\|W\|_* + \frac{1}{2}\|W - (\hat{X}^{k+1,i} + \hat{\gamma}^{k,i})\|_F \},$$

$$\hat{\gamma}^{k+1} = \hat{\gamma} + (\hat{X}^{k+1,i} - \hat{Z}^{k+1,i}).$$

In a dense network, the number of transceivers (Tx or Rx) can be larger than 2. When all transceivers need a relay to forward data, how the UAV find the optimal location? An example of four transceivers is shown in Fig. 3, in which A wants to connect C and B wants to connect D. Nevertheless, they cannot transmit data directly due to mmWave’s limited range and relays are needed to help the forwarding. The naive solution is to utilize 2 UAVs to build these 2 links. But such a solution wastes the UAV resources, especially when the number of transceivers is further increased.

The enhanced AutoRelay adopts only one UAV to serve multiple transceivers in dense networks. The UAV operates the same procedure of core AutoRelay with only minor modifications of Step 1 and 2. In Step 1, the link qualities of multiple transceivers should be considered into the matrix $Q$. We redefine the matrix as $Q(i,j,k) = R_A R_B R_C R_D$, where $R_A$ is the link quality between the position of $R_A$ and
First, one certain UA $V$ takes a picture of all transceivers, and find their locations accurately? Then, how AutoRelay drives multiple relays collaboratively?

**A. Environment Setting**

In the simulations, we create a 3D space scene which simulates the real life scene. Specifically, in the 3D space based on the point $(0, 0, 0)$ and $c(n_1, n_2, n_3)$, we put a cuboid representing the building in the environment marked by $b_1, b_2, ..., b_8$ based on the value of $n_1, n_2, n_3$. The settings of these points are listed in the Table II.

Since the signal attenuation through the obstruction is severe, the optimal location for mobile relay will not lie in the obstructed path. Thus, these points are ruled out when finding the optimal location for best link quality. To simulate the real life scene, in addition to considering the large-scale path loss, we add a random value with normal distribution to obtain the experiment data. Additionally, the values in the matrix have all been normalized by dividing the value by the maximum value in the matrix.

In the simulation, we first consider the base methods to achieve the optimal location. We traverse the whole real quality matrix to find the real optimal location with the best link quality. Afterwards, we compare the performance of AutoRelay with that of the classic algorithm, K-Nearest Neighbor (KNN), and that of the state-of-the-art algorithm Tensor Recovery (TR). In KNN, during the flying process to the optimal location, the relay updates the values of the 6 nearest neighbors of the current location and finds the new optimal location until reaching the final optimal location. While in TR, the link quality matrix is recovered once with the sampled data then the optimal location is determined.

To evaluate the final optimal location quality, we conducted $T$ trials and in each trial, $N$ simulations were done and we measure the performance $E$ of each trial by

$$E = \sum_{i=1}^{N} (Q_i - Qr)^2,$$

where $Qr$ is the link quality of the real optimal location. The metric $E$ evaluates the difference between the achieved results and the real optimal location. Better algorithms have smaller value since the achieved results are closer to the link quality of the real optimal location. We calculate the average of the trial performances $E$ to compare the general performance of the three algorithms.

**B. AutoRelay With Multiple Relays**

On the contrary to dense networks, one relay may be not enough to support a connection between transmitters and receivers in sparse networks due to the limited range of one hop. As shown in Fig. 4, since the distance between AB and CD is too long, at least two mobile relays are needed. Then, how AutoRelay drives multiple relays collaboratively and find their locations accurately?

The enhanced AutoRelay for multiple relays is as follows: First, one certain UAV takes a picture of all transceivers, estimates the distance among them, and determine how many relays are needed. Second, multiple relays depart from the same place. Thus, these relays ($R_1$ and $R_2$ in Fig. 4) are connected at the beginning. Third, AutoRelay sets the constraint that relays should maintain their connection during their movement. Fourth, in Step 2, the optimal relay locations are selected according to the metric $\max(R_{A\leftarrow B} R_{B\leftarrow C} R_{C\leftarrow D})$.

**C. AutoRelay for Directional Beams**

Omni-directional broadcast is not the unique manner for mmWave. Directional transmission is also usual in order to increase the communication range and reduce the interference. Different from omni-directional broadcast, mmWave transmitter could adjust its beam’s direction to partially change the link quality of one given location.

As shown in Fig. 5, we design a beam adjustment mechanism in enhanced AutoRelay to improve the link quality in directional transmission. This mechanism requires the transmitter to rotate its beam and scan the optimal direction to the relay. The beam adjustment may trigger the movement of relay due to the online update feature of AutoRelay. The adjustment and movement will iterate and converge quickly because mmWave’s antenna array can control a beam into a very small angle, e.g., smaller than 7 degree.

![Fig. 4. Example of AutoRelay for multiple transceivers and multiple relays.](image1)

![Fig. 5. Example of AutoRelay adjustment for directional beam.](image2)

**TABLE II**

<table>
<thead>
<tr>
<th>Obstruction Points Setting in the Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$ $(0, \frac{1}{2}n_2, 0)$</td>
</tr>
<tr>
<td>$b_3$ $(\frac{3}{4}n_1, n_2, 0)$</td>
</tr>
<tr>
<td>$b_5$ $(0, \frac{1}{2}n_2, n_3)$</td>
</tr>
<tr>
<td>$b_7$ $(\frac{3}{4}n_1, n_2, n_3)$</td>
</tr>
</tbody>
</table>

**VI. Performance Evaluation**

In this section, we conduct extensive simulations to evaluate the performance of AutoRelay by comparing with the classic and the state-of-the-art algorithms.
In addition, to compare the stability of the three algorithms, the standard error is calculated by

$$Error_s = \sqrt{\frac{\sum_{i=1}^{N} (Q_i - \bar{Q})^2}{N - 1}},$$

where $\bar{Q} = \frac{\sum_{i=1}^{N} Q_i}{N}$. The metric $Error_s$ evaluates the standard error of the results of different simulations. The smaller the $Error_s$ is, the more stable results the algorithm produces.

In order to measure the efficiency, the time costs of the three algorithms are also compared. We evaluate the time cost for each trial by the average time cost of each simulation and the formulation is given by

$$\bar{t} = \frac{\sum_{i=1}^{N} t_i}{N}.$$  

Experiments are conducted with different data sizes to acquire more results about the performance.

The procedures of simulation are as follows:
- First, initialize with the parameters of the original data. Setup the 3D space, mark the point $o, c, s$ and calculate the initial quality matrix $Q_o$.
- Second, simulate the relay’s flying process. AutoRelay, KNN and TR update the link quality matrix respectively and record the link quality of the final location for relay.
- Finally, repeat the simulations several times with different parameters and compare the performances among three algorithms.

### B. Simulation Results

In Fig. 6, we plot the performance comparisons among three algorithms on multiple trials with different matrix sizes.

To evaluate the quality performance, it can be seen from (a), (b) and (c) that, in general, AutoRelay outperforms KNN and TR. According to the measurement of link quality given by (17), we find that the less the performance $E$ is, the better the link quality is acquired. Specifically, in (a), AutoRelay achieves the performance value $0.412 \times 10^{-3}$, KNN achieves $1.289 \times 10^{-3}$ and TR achieves $0.703 \times 10^{-3}$. Comparing the results achieved by the three algorithms, we find that AutoRelay is about 3 times better than KNN and 2 times better than TR. In (c), where the performance of TR is close to AutoRelay whose values are $0.586 \times 10^{-3}$ and $0.430 \times 10^{-3}$, respectively, AutoRelay is about 5 times better than KNN whose value is $1.935 \times 10^{-3}$. From (b), AutoRelay also has a smaller value of $E$ indicating better performance than TR and KNN. In total, it can be seen from the first three figures that performance of AutoRelay is better than KNN and TR.

Moreover, AutoRelay has the highest stability. Specifically, in (a), the values of $Error_s$ for each algorithm are $0.189 \times 10^{-3}$ for AutoRelay, $0.735 \times 10^{-3}$ for TR and $0.822 \times 10^{-3}$ for KNN. AutoRelay produces less fluctuation in achieved quality performances than TR and KNN. While in (b), the corresponding values of $Error_s$ are $0.238 \times 10^{-3}$ for AutoRelay, $0.441 \times 10^{-3}$ for TR and $0.635 \times 10^{-3}$ for KNN, from which we can easily conclude that the results produced by AutoRelay have relatively smaller standard error and thus AutoRelay is more stable. In (c), similarly, AutoRelay gives a more stable results than TR and KNN which have relatively more fluctuation in the produced results. With different data sizes, AutoRelay also provides a more stable performance than KNN and TR. When measuring the stability of the algorithms, AutoRelay outperforms KNN and TR.

Fig. 6(d)(e)(f) show the time cost of three algorithms. From (d), we find that the average time cost of AutoRelay is 79...
while KNN and TR are around 35. From (e) and (f), it can be seen that the time cost of AutoRelay is about 2 times the time costs of KNN and TR. In general, AutoRelay has relatively more time cost than KNN and TR. However, the transmission can be established before reaching the optimal location by AutoRelay. That is, shortly after taking off from the starting point, when reaching the first-round relay location, the relay starts to transmit data from the source to the destination. Under such circumstances, the actual transmission begins much earlier than reaching the final optimal location, which makes up the disadvantage of AutoRelay in time cost.

In summary, AutoRelay outperforms KNN and TR in acquiring better quality and achieving higher stability, but with more time cost.

VII. RELATED WORK

Relays are common solutions to maintain the wireless link for disconnected transmitter and receiver. Plenty of works have contributed to the relay research community. We classify existing works into two categories: static and mobile relays.

A. Static Relays

At the beginning, static relays attract the research and trigger the real applications. Static relays can long-term work because of its stable power supply. However, a large amount of redundant relays are need to increase the connectivity of networks. From theoretical direction, cooperative strategies and capacity theorems are analyzed in [13]. From application direction, generalized signal alignment in [15] is investigated for multi-user MIMO two-way relay channels and its degrees of freedom is discussed in [21] with clustered pairwise exchange.

B. Mobile Relays

With the development of UAVs, lots of works have contributed to the adoption of mobile relays. Pinkney et al. [18] propose the drone based battlefield information system. Drones are used as platforms for a high capacity trunk radio relay to form a battlefield broadcast system. Then, UAV relays are widely studied in performance enhancement. For example, Zhan et al. [25] comprehensively analyze the performance of UAV based wireless relay communications and propose specialized optimization methods to enhance the performance. Using mobile relays to offer high-quality LTE services is discussed in [4]. Key techniques, such as the group mobility and the local service support are applied to improve the efficiency.

Energy efficiency is another useful purpose in mobile relays. In order to prolong the lifetime [22], mobile relays are in charge of most data collection tasks from sensor nodes to sink. In [14], cooperative movement strategy is applied in a team of UAVs to form an energy-efficient relay networks.

However, these solutions cannot exactly fit the mobile relay problem for mmWave communications because they neither consider the mmWave propagation characteristics nor ensure the quick search of the optimal location. Thus, we need to develop a new mobile relay method specialized for mmWave.

VIII. CONCLUSION

Considering the disconnection between mmWave transmitter and receiver, we propose to use mobile relays to bypass the obstacles or prolong the communication range. To tackle the challenges of accurate location determination and mobile transceivers, we design the AutoRelay method to sample real-time link quality and keep updating the matrix to quickly determine the optimal location. Simulation results show that the AutoRelay algorithm finds more accurate relay location and produces more stable results than existing methods.

We believe mobile relays has wider implications for mmWave communication design than explored in this paper. First, more environmental factors can be considered to improve the accuracy and stability. Cooperative relay strategy is one of our future works.

REFERENCES


Linghe Kong (S’09–M’13) received the B.E. degree from Xidian University in 2005, the master’s degree from TELECOM SudParis in 2007, and the Ph.D. degree from Shanghai Jiao Tong University in 2012. He was a Post-Doctoral Researcher with Columbia University, McGill University, and the Singapore University of Technology and Design. He is currently a Research Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University. His research interests include wireless communication, sensor networks, mobile computing, Internet of things, and smart energy systems.

Linsheng Ye received the B.E. degree from Shanghai Jiao Tong University in 2016, where he is currently pursuing the master’s degree with the Department of Computer Science and Engineering. His research interests include wireless communication and data center.

Fan Wu received the B.S. degree in computer science from Nanjing University in 2004, and the Ph.D. degree in computer science and engineering from the State University of New York at Buffalo in 2009. He has visited the University of Illinois at Urbana–Champaign as a Post-Doctoral Research Associate. He is currently a Full Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. His research interests include wireless networking and mobile computing, algorithmic network economics, and privacy preservation.

Meixia Tao received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology in 2003. She was an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, from 2004 to 2007. She is currently a Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. Her current research interests include caching and multicasting in content-centric wireless networks, resource allocation, interference management and coordination, MIMO techniques, and physical layer security.

Guohai Chen received the B.S. degree from Nanjing University in 1984, the M.E. degree from Southeast University in 1987, and the Ph.D. degree from The University of Hong Kong in 1997. He is currently a Distinguished Professor with Shanghai Jiao Tong University, China. He had been invited as a Visiting Professor by many universities, including the Kyushu Institute of Technology, Japan, in 1998, The University of Queensland, Australia, in 2000, and Wayne State University, USA, from 2001 to 2003. He has a wide range of research interests with focus on sensor network, peer-to-peer computing, high performance computer architecture, and combinatorics.

Athanasios V. Vasilakos (M’00–SM’11) is currently a Professor with the Luleå University of Technology, Sweden. He is the General Chair of the European Alliances for Innovation. He was/is an Editor for many technical journals, such as the IEEE Transactions on Network and Service Management, the IEEE Transactions on Cloud Computing, the IEEE Transactions on Information Forensics and Security, the IEEE Transactions on Cybernetics, the IEEE Transactions on NanoBioscience, the IEEE Transactions on Information Technology in Biomedicine, the ACM Transactions on Autonomous and Adaptive Systems, and the IEEE Journal on Selected Areas in Communications.