Research Article

Performance Study and Optimization of 3D-MANET: A New Analytical Perspective Based on Zipf’s Law

Zairan Cheng and Ying Liu

1School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
2Signal and Communication Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing, China
3Weihai International College, Beijing Jiaotong University, Weihai, China

Correspondence should be addressed to Ying Liu; liuying@bjtu.edu.cn

Received 6 January 2022; Revised 15 June 2022; Accepted 23 June 2022; Published 8 September 2022

Academic Editor: Danfeng Hong

Copyright © 2022 Zairan Cheng and Ying Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies the throughput capacity and delay scaling laws in a three-dimensional mobile ad hoc network (3D-MANET) under different routing schemes. Previous work generally assumed that nodes follow a uniform distribution or a power-law distribution to move in the network. From the perspective of the entire network, it is difficult for this network model to reflect communication entities’ distribution in real 3D space. Moreover, the research results of analyzing network performance using different routing schemes are limited, and the research work is insufficient. With formerly related studies different, we propose a cell-gridded network model that considers the actual environment with cells of the node aggregation degree, which follows Zipf’s law with exponent γ. And our model can cover a variety of distribution scenarios with changes in the γ value. The packet delivery rate, network capacity, and delay performance of 3D-MANET adopting the traditional two-hop nonredundant and redundant relay routing scheme are examined utilizing theoretical tools such as probability theory, random process, and queuing theory. We propose a wireless access point (WAP-) enabled multihop relay routing scheme. By deploying WAP in cells with a high γ, nodes can access WAP and broadcast packets, which accelerates the delivery of packets, and the results obtained by applying this scheme indicate that compared with the two-hop relay scheme, the WAP multihop relay effectively improves the delay performance and the transmission efficiency with less loss of capacity performance. Additionally, a better delay-capacity trade-off performance is achieved. Finally, we discuss the influence of parameters such as the number of network nodes n, the number of network cells m, the redundancy r, and γ on the capacity and delay. The analysis results confirm that exploiting the users’ distribution status information and dividing the cells reasonably will save deployment costs and further improve network performance.

1. Introduction

In a mobile ad hoc network (MANET), network nodes (e.g., users’ mobile devices, sensors, cars, aircraft, fleets, and submarines) cooperate in providing network services that are traditionally implemented by network infrastructure (such as routers, switches, and servers). IEEE 802.11 ad hoc mode enables these facilities with a wireless interface to activate a communication session with other devices and finish data transmission without network infrastructure [1]. This type of network can be rapidly deployed and configured and possesses self-organizational solid, robust, and invulnerable qualities. Numerous network paradigms developed from MANET, including wireless mesh network [2], wireless sensor network [3], opportunistic network [4], and vehicle ad hoc network [5], have demonstrated exceptional effectiveness in various application scenarios.

The application scope of the MANET has been extended to three-dimensional (3D) space and derived the flying ad hoc network [6], aeronautical ad hoc network [7], 3D-vehicular ad hoc network [8], and other emerging 3D-MANET paradigms [9] with the development of aeronautical communication and the rise of unmanned aerial vehicles (UAV).
Similarly, 3D-MANET can be rapidly deployed and flexibly configured, which has irreplaceable advantages for various critical scenarios. For example, it will ensure the regular communication of military units such as control centers, aircraft, UAVs, troops, and warships, assists the multi-UAV system in completing the environmental monitoring work, and provides emergency communications in the disaster areas to aid in the completion of rescue missions, and it can also connect underwater vehicles and massive sensors to support marine monitoring. Besides, it will also play a key role in deep space communications and interstellar exploration [7, 10–17].

For the foreseeable future, the implementation field of 3D-MANET will become more extensive. For instance, with the new wave of urbanization, the urban environment will see an increase in 3D-VANET scenes generated by overpasses, viaducts, tunnels, slopes, and other structures. Because of this, it is critical to explore fundamental concerns like network throughput capacity (hereinafter referred to as capacity) and delay performance, as well as the impact of various parameters on network performance, which can provide a theoretical reference for the proper structure and routing design of 3D-MANET.

1.1. Related Work. Network performance evaluation has always been an open-ended primary theoretical topic. Gupta and Kumar devised a framework for analyzing the performance of large-scale wireless networks in the early 2000s [18, 19], and the current work on the performance of 3D-MANET also followed the main ideas of this framework.

Li et al. [20] investigated the capacity bounds of 3D-MANET in the state of uniform and nonuniform distribution of nodes by employing Shannon’s theorem [21]. In nonuniform networks, the distribution of nodes conforms to the general nonhomogeneous Poisson process. Compared with the results presented by Gupta and Kumar in literature [19], the results of Li et al. indicate that under the physical interference model, the lower limit of the capacity of the two types of networks depends on the power propagation environment, that is, the path loss index. For networks in these two dimensions, when the path loss index is in a particular range, 3D-MANET may have a lower network throughput capacity than 2D-MANET.

Bai et al. [22] considered the nodes distributed according to shot noise Cox process [23] and obtained the achievable throughput capacity of 3D inhomogeneous wireless networks using redesigned spheroidal percolation model [24]. The “information pipes” between clusters are established through the clustering operation of network nodes, and the lower bound of the network capacity between clusters is deduced.

Jeong and Shin [25] took the 3D erasure networks with the random distribution of nodes as the research object, introduced a routing protocol using a percolation highway in 3D space, and utilized exponential and polynomial power-law path loss models to simulate the erasure probability of packet transmission. They used percolation theory and the maximum-flow minimum-cut theorem to determine each path loss model’s throughput capacity scaling law.

Wei et al. [26] studied the network capacity performance of three-dimensional cognitive radio networks (3D-CRNs) and evaluated the effect of path loss index on the capacity of 3D-CRNs. Since then, they have investigated the asymptotic capacity performance of 3D wireless social networks in [10] and proposed the concept of the social contact group. The members of the social contact group open communication sessions at random, but the location information of the members frequently varies, and the communication link between members also changes. This type of analysis may not accurately reflect the most realistic network performance.

Wang et al. [27, 28] considered a scenario in which network nodes were uniformly distributed, and the routing strategy adopted a two-hop relay routing algorithm with redundant packets. Use Markov chain theory [29] to describe the packet delivery process, and then obtain the closed-form expression for 3D-MANET capacity and packet delivery delay. Although the authors derive theoretical results such as network capacity, expected delay value, and delay relative standard deviation in detail, they do not develop a delay-capacity trade-off relationship.

Guo et al. [30] considered a high-density urban area communication scenario to study the channel capacity of 3D mobile UAV networks, which followed stochastic geometry theory and MANET design principles and designed a 3D Poisson point process mobile UAV network model that permits drones to move in 3D space freely. The ground user equipment adopts the nearest neighbor association strategy and the random association strategy to select the UAV as the serving base station. The rest are interfering base stations. Then, through the Laplace transform of the distance distribution between the UAV and the user and the interference of other base stations, it is concluded that the downlink network coverage is proportional to the channel capacity under the two association strategies.

Muneeswari and Manikandan [31] suggested a new 3D-MANET routing scheme based on the reinforcement learning algorithm. The indicators considered in this scheme are delay between nodes, the distance between nodes, node stability, node mobility, and node connectivity. Using NS3.26 [32], an extensive simulation analysis of the routing protocol in packet delivery rate, end-to-end delay, routing packet overhead, energy consumption, and security performance is carried out in 3D space.

1.2. Motivations and Contributions. Based on a summary of these prior related research works, we believe three issues require urgent resolution.

(i) As stated above, it is difficult to depict the distribution of network communication entities in the real 3D space by the network mobility model adopted by the existing theoretical results to facilitate the model solution. Whether it is a uniform or nonuniform moving model, there are discrepancies between the actual application scenarios, and the results obtained are difficult to extend to the 3D space directly. In addition, the 2D network model...
directly ignores the influence of the nonlight of sight (NLOS) transmission, making the network model abstracted as 2D topology too ideal. In contrast, a 3D space with a specific height can ensure the light of sight (LOS) transmission, and the influence of NLOS transmission can be approximately ignored, so directly in 3D modeling in space tends to produce more accurate results.

(ii) There are limited research findings on analyzing network performance utilizing diverse routing algorithms. Especially for 3D space, the number of nodes aggregated in distinct space areas varies, so the number of source-destination node pairs, relay nodes, and destination nodes contained in different space areas is different. Consequently, each area’s communication transmission situation and transmission probability are distinct.

(iii) Previous studies did not explore the trade-off between 3D-MANET optimal network delay and capacity. Some studies’ results only discussed capacity without considering delay or only discussed delay without paying attention to capacity, which is unrealistic.

The majority of our efforts have been devoted to solving these issues in a targeted manner. The following are the main contributions of this paper:

(i) We propose a novel 3D-MANET of a cell-gridded network model based on Zipf’s law. The model can simulate many node distribution scenarios in 3D space by changing the parameter \( \gamma \).

(ii) Then, the packet delivery rate, network capacity, and delay performance of 3D-MANET are deduced and analyzed in detail, and the setting value of optimal packet redundancy and its related optimal delay performance is given.

(iii) We propose a WAP-enabled multihop relay routing scheme to optimize network performance, which has proven to improve latency performance and achieve a better delay-capacity trade-off performance.

(iv) Finally, we discuss the impact of each parameter on the network performance.

2. System Model and Framework

In this section, we detail the preparations for 3D-MANET capacity and delay analysis, which includes a cell-gridded 3D-MANET network (mobility) model based on Zipf’s law, an interference model, traffic pattern, related model assumptions, and selected routing scheme.

For the purpose of getting the generalization results, this paper discusses a large-scale multihop 3D-MANET made up of any communication entity found on land, sea, or air. If the network needs to send data through the Internet, it will designate some nodes as gateways to ensure connectivity with the Internet.

2.1. Network and Mobility Model. Previous studies usually assume that nodes move in a uniform distribution in the network, but not all nodes move in these ways in actual scenarios. On the other hand, although some power-law distribution models can better describe the nonuniform distribution movement characteristics of some nodes, from the perspective of the whole network, these nodes are still roughly uniformly distributed. For prediction-type models like Gauss-Markov mobility models, they mainly simulate the motion behavior of nodes, such as speed and direction. In view of this, to propose a mobile model suitable for 3D-MANET, this paper takes the actual space scene as the breakthrough point, and based on the large amount of research data presented in literature [33, 34] Schläpfer et al. and Pappalardo et al., we discovered that in the multitype mobile scenes, the number of nodes aggregated in different spatial locations is different, and the distribution of nodes in space roughly follows Zipf’s law.

Inspired by this research work and experimental data, we consider a cubic space network with \( n \) nodes that is partitioned into \( m \) equal-sized unit cubic cells, as indicated in Figure 1. Time is segmented into time slots of equal length. At the start of each time slot, the node moves from one cell to another, staying in that cell for the remainder of the time slot. According to Zipf’s law [33, 35], for any node in the network, the probability of moving to the cell \( j \) is

\[
p_j = F(m) \cdot j^{-\gamma} \quad 1 \leq j \leq m,
\]

where \( \gamma \) is defined as the distribution exponent of the density of nodes in cell \( j \). According to Zipf’s law, \( j \) also represents the aggregation degree of nodes in the cell. The smaller \( j \) is, the more nodes are aggregated in the cell, and the greater the cell’s popularity. By the \( \sum_{j=1}^{m} p_j = 1 \) and the Maclaurin-Cauchy integral test method [36], the \( F(m) \) in order sense (Given two functions \( f(n) \) and \( g(n) \), we have the following notations: \( f(n) = \Theta(g(n)) \) means \( f(n) \) is asymptotically tight-bounded by \( g(n) \), i.e., \( \lim_{n \to \infty} |f(n)/g(n)| = c \), where \( c \) is a positive constant,) can be obtained:

\[
F(m) = \left( \sum_{j=1}^{m} j^{-\gamma} \right)^{-1} = \left( \int_{1}^{m} j^{-\gamma} \right)^{-1} = \begin{cases} \Theta(m^{\gamma-1}), & 0 < \gamma < 1, \\ \Theta \left( \frac{1}{\log m} \right), & \gamma = 1, \\ \Theta(1), & \gamma > 1. \end{cases}
\]

We can observe that a few cells have a greater density of nodes, whereas the bulk of cells has a sparse node distribution. Imagine that frequently dense concentration of network users in specific hotspot locations in the cities where we live, while users are sparsely spread in other areas. It should also be noted that the nodes included within these colored cells still adhere to the Zipf distribution in the height dimension. For demonstration purposes, only the bottom surface of the cell is colored; the deeper the color of the cell indicates, the more nodes are clustered [37].
2.2. Interference Model and Transmission Pattern. This paper assumes that the transmission between nodes in the same time slot satisfies the protocol interference model (PrIM) constraint proposed by Gupta and Kumar [18]. If node $N_i$ transmits data to node $N_j$ through a channel, when

$$|N_i - N_j| = (1 + \Delta)|N_i - N_j|,$$  \hspace{1cm} \text{(3)}

then $N_j$ can successfully receive data, where the identifier $N$ also represents the location of the node, $|N_i - N_j|$ represents the euclidean distance between nodes $i$, $j$, $N_{pr}$ represents any other node transmitting across the same channel, and the coefficient $\Delta (\Delta > 0)$ simulates a protective space to prevent the transmission interference of nearby nodes on the same channel.

According to the PrIM, to enable multiple links to transmit simultaneously without interfering with each other, we borrow the scheduling transmission scheme [38] of Liu et al. Considering a simultaneous transmission group (STG), in which any two cells in a network can sequentially get transmission opportunities in $\alpha^3$ time slots based on the time division multiple access technologies. Meanwhile, we utilize the unicast transmission mode described by the sequential transmission model. That is, the nodes in the network are numbered and indexed from 1 to $n$, and then, the established 3D-MANET network model has $n/2$ distinct source-destination node pairs, which are independent of each other. Each node also serves as a relay node, assisting other $(n - 2)$ nodes in forwarding packets.

2.3. Two-Hop Relay Routing Algorithm. Since the two-hop relay routing scheme was first used by Grossglauer and Tse in literature [39], due to the scheme’s simplicity and efficiency, as well as the fact that it is conducive to the analysis and study of capacity and delay performance, this scheme and its extensions have become a trendy routing scheme for MANET networks. Research teams mostly use this algorithm to evaluate network performance. This paper will analyze and compare the network performance using the nonredundant and redundant relay routing methods for the two transmission modes. Note that the network topology of 3D-MANET is highly dynamic, and nodes in the network can freely move to any position in the 3D space. Therefore, this paper does not use the ACK message confirmation mechanism to ensure that the data packet successfully reaches its destination node. This paper assumes that the destination node can successfully receive the data packet if it is within the source node’s transmission range and unaffected.

2.3.1. The Two-Hop Nonredundant Relay. As illustrated in Figure 4, the two-hop non-redundant relay routing algorithm completes the delivery of data packets in two ways. One route is that the source node $S$ transmits the packet directly to the destination node $D$, and the other is that the source node $S$ transmits the packet via the relay node $R$.

2.3.2. The Two-Hop Redundant Relay. Redundant relay means data packet redundancy, which means multiple copies of a data packet are delivered to the destination node across multiple relay nodes in the network. This technology effectively increases the transmission opportunities of data packets in the network, which can reduce the time required for the data packet to reach the destination node. Based on the nonredundant relay algorithm, this paper selects this routing algorithm to analyze the network performance to get a comparative conclusion and discuss the influence of redundancy on the network performance.

As depicted in Figure 5, there are also two ways for the redundant relay routing algorithm to complete the delivery of data packets. The difference from the nonredundant relay is that the source node $S$ will transmit to $r$ relay nodes at most $r$ copies of the data packet (that is, the packet redundancy is $r$), and the destination node $D$ may successfully receive the data packet from either source node $S$ or one of the relay nodes $R'$. Therefore, there are no more than $r + 1$ copies of a particular packet in the network.
3. Capacity and Delay Performance

Analysis of 3D-MANET

Based on the preliminary work in the preceding section, this section will detail the packet delivery rate, capacity, and delay performance of 3D-MANET using the two-hop nonredundant relay and redundant relay routing scheme. Additionally, we will discuss optimal packet redundancy and its corresponding optimal latency performance. First, we give some relevant definitions.

3.1. Definitions. In terms of network capacity, MANET is different from the mainstream cellular network in its definition method. The capacity of MANET is characterized by the amount of data successfully transmitted per second on a single node or the entire network [18, 19]. In terms of
delay, the definition of end-to-end delay in the field of computer networks is utilized. This paper defines the capacity and delay as follows to facilitate understanding and analysis of performance:

**Definition 1** (per-node (throughput) capacity). The per-node capacity is the maximum rate $\lambda$ (unit is packets/time slot) at which the network can sustain a stable state of transmission.

**Definition 2** (network (throughput) capacity). The network capacity ($C$) is defined as the sum of the per-node capacity of all the nodes in the network.

**Definition 3** (end-to-end delay). Normally, the end-to-end delay of a data packet is the time required for the packet to transit from the source to the destination. It comprises two components: the packet’s transmission delay and the packet’s queuing delay at the source node, abbreviated as the delay in the following. The average delay of the network is defined as the expectation obtained by averaging the delays of all packets and source-destination node pairs over a lengthy period of time, and the unit is the time slot.

3.2. Two-Hop Nonredundant Relay Routing Scheme’s Network Capacity. The theoretical results of Neely and Modiano, Chen et al. [40–42] show that if the nodes in the network satisfy the i.i.d mobility model, then the network capacity depends on the node distribution state in the steady-state of the network. We consider at least one packet delivery event to occur in a given time slot when there are at least two nodes in the same cell, depending on the transmission pattern. Moreover, this study assumes that data packets arrive with the Bernoulli process, meaning that either a single data packet arrives or none
arrives within the current time slot. We assume that all nodes in the network receive packets at the same rate (i.e., \( \lambda_i = \lambda \) for all \( i \in N_s \{|1 \leq i \leq n\} \)). Let \( p_{\text{sum}} \) represent the average probability that a cell contains at least two nodes (i.e., the packet delivery rate), and \( p_{\text{rd}} \) indicates the average probability of at least one pair of source-destination nodes in the same cell (that is, the probability that the packet is directly delivered from the source node to the destination node without a relay node). Then, for each time slot, the total amount of data packets generated in the network is \( n \lambda \), and the maximum amount of data packets delivered in the network is \( m p_{\text{sum}} \) (\( m p_{\text{sum}} \) implies the sum of cells contain at least two nodes). Respectively, the maximum amount of packets delivered directly by the single-hop transmission of source-destination node pairs is \( m p_{\text{rd}} \) (\( m p_{\text{rd}} \) represents the sum of cells with at least one source-destination node pair). In each cell with at least two nodes, all nodes in the cell have the same probability of being randomly selected to be the sending node or the receiving node, and the sending node triggers the “source node → relay node” and “relay node → destination node” with equal probability of these two transmission modes. Therefore, the maximum amount of packets transmitted by two-hop communication in the network through source node-relay node-destination node is \((1/2)m (p_{\text{sum}} - p_{\text{rd}})\) for each time slot [37]. We can get the relational expression as follows:

\[
C = n \lambda \leq m \bar{p}_{\text{rd}} + \frac{1}{2} m (\bar{p}_{\text{sum}} - \bar{p}_{\text{rd}}),
\]

i.e.,

\[
C = n \lambda \leq \frac{1}{2} m (\bar{p}_{\text{sum}} + \bar{p}_{\text{rd}}).
\]

Next, we use Zipf’s law to derive the relevant probabilities \( p_{\text{sum}} \) and \( p_{\text{rd}} \) and then get the result of the network capacity.

According to the network mobility model described in Section 2.1, let \( p_{\text{sum}}^j \) denote the probability of the event that cell \( j \) has at least two nodes; we have

\[
p_{\text{sum}}^j = \frac{1}{\alpha^3} \left[ \left( \begin{array}{c} n \\ 0 \end{array} \right) (1-p_j)^n - \left( \begin{array}{c} n \\ 1 \end{array} \right) p_j (1-p_j)^{n-1} \right],
\]

where \( \left( \begin{array}{c} n \\ 0 \end{array} \right) (1-p_j)^n \) and \( \left( \begin{array}{c} n \\ 1 \end{array} \right) p_j (1-p_j)^{n-1} \) represent the probability that there is no node in cell \( j \) and the probability of a single node, respectively.

Combine formulas (1) and (2) to simplify formula (6). Since this paper considers a large-scale network, the important limit formula can be used to write formula (6) as

\[
p_{\text{sum}}^j = \frac{1}{\alpha^3} \left[ 1 - \left( \begin{array}{c} n \\ 0 \end{array} \right) (1-p_j)^n - \left( \begin{array}{c} n \\ 1 \end{array} \right) p_j (1-p_j)^{n-1} \right] = \Theta \left( \frac{1}{\alpha^3} \right) (1 - \frac{1}{\alpha^3})
\]

\[
\quad = \Theta \left( \frac{1 - (1-p_j)^n}{\alpha^3} \right)
\]

\[
\quad = \Theta \left( \frac{1 - (1-p_j)^{nP_j}}{\alpha^3} \right) = \Theta \left( \frac{1 - e^{-\alpha p_j}}{\alpha^3} \right).
\]

(7)

Furthermore, since \( N_p \) represents the number of nodes in the cell \( j \), when \( n \) is large, \( (-np_j) \) can be interpreted as the number of nodes in the cell \( j \) approaching zero; formula (7) is further expressed as

\[
p_{\text{sum}}^j = \Theta \left( \frac{np_j}{\alpha^3} \right).
\]

(8)

Given that the value of \( p_j \) depends on \( \gamma \), it is necessary to analyze the situation in which \( \gamma \) takes different ranges of values. Next, we utilize the Maclaurin-Cauchy integral test method to determine the accurate values of the aforementioned packet delivery rate \( p_{\text{sum}} \) when \( \gamma \) assumes different values.

Case 1. \( 0 < \gamma < 1 \), \( p_j = \Theta(m^{-\gamma}). \) and

\[
\bar{p}_{\text{sum}} = \frac{1}{m} \sum_{j=1}^{m} p_{\text{sum}}^j = \frac{1}{m} \sum_{j=1}^{m} \Theta \left( \frac{np_j}{\alpha^3} \right) = \frac{1}{m} \sum_{j=1}^{m} \Theta \left( \frac{n \cdot m^{-\gamma} \cdot j^{-\gamma}}{\alpha^3} \right)
\]

\[
\quad = \frac{1}{m^2} \Theta \left( \frac{n \gamma}{\alpha^3} \right) \sum_{j=1}^{m} \left( \frac{m \gamma}{\alpha^3} \right) j^\gamma = \Theta \left( \frac{n \gamma}{\alpha^3} \right) \left( \frac{1 - m^{-\gamma}}{1 - \gamma} \right).
\]

(9)

Similarly, we can also determine the value when \( \gamma = 1 \) and \( \gamma > 1 \).

Case 2. \( \gamma = 1 \), \( p_j = \Theta \left( \frac{1}{(j \log m)} \right) \), and \( \bar{p}_{\text{sum}} = \Theta \left( \frac{n}{m \log m} \right) \).

Case 3. \( \gamma > 1 \), \( p_j = \Theta(\gamma^{-j}) \), and

\[
p_{\text{sum}}^j = \Theta \left( \frac{n}{m \gamma^{-1}} \cdot \frac{1}{\gamma - 1} \right).
\]

(10)

The order meaning of the packet delivery rate \( \bar{p}_{\text{sum}} \) is

\[
\bar{p}_{\text{sum}} = \begin{cases} 
\Theta(\beta), & 0 < \gamma < 1, \\
\Theta(\beta \cdot \frac{1 - m^{-\gamma}}{1 - \gamma}), & \gamma = 1, \\
\Theta(\frac{\beta}{\gamma - 1}), & \gamma > 1,
\end{cases}
\]

where \( \beta = \frac{n}{m \gamma} \).
Similar to the idea of deriving the value of the packet delivery rate \( \rho_{\text{sum}} \), we will proceed to analyze the value of \( \rho_{sd} \). Let \( p_{sd}^j \) denote the probability of the event that the number of source-destination node pairs in cell \( j \) is not less than one. According to the transmission mode assumed in Section 2.2, the 3D-MANET network model established in this paper contains \( n/2 \) independent source-destination node pairs, so the probability that any of the nodes which constitute source-destination node pairs is not in the cell \( j \) is \((1 - p_j^1)^{n/2}\); then, the probability that the cell \( j \) does not have any source-destination node pair is \((1 - p_j^1)^{n/2}\); thus

\[
p_{sd}^j = \frac{1}{\alpha^3} \left[ 1 - \left( 1 - p_j^1 \right)^{n/2} \right].
\]

(12)

The simplification process of \( \rho_{sd}^j \) and \( \rho_{sum}^- \) is similar. The specific derivation process is as follows:

\[
p_{sd}^j = \frac{1}{\alpha^3} \left[ 1 - \left( 1 - p_j^1 \right)^{n/2} \right] = \Theta \left( \frac{1 - \left( 1 - p_j^1 \right)^{(1/2)} \cdot (n/2) p_j^1}{\alpha^3} \right)
\]

\[= \Theta \left( 1 - e^{-\alpha p_j^1} \right) = \Theta \left( \frac{n}{2\alpha} \cdot \alpha^2 p_j^1 \right).
\]

(13)

In order to determine the exact value of the probability \( \rho_{sd}^j \) that the aforementioned data packet is directly delivered from the source node to the destination node, it is necessary to analyze the situation where \( \gamma \) takes different ranges of values because \( p_j \) is depended on \( \gamma \).

Case 1. \( 0 < \gamma < 1 \), \( p_j = \Theta(m^{\gamma-1} \cdot j^{-\gamma}) \), and

\[
\rho_{sd} = \frac{1}{m} \sum_{j=1}^{m} p_{sd}^j = \frac{1}{m} \sum_{j=1}^{m} \Theta \left( \frac{n}{2\alpha} \cdot \alpha^2 p_j^1 \right)
\]

\[= \frac{1}{m} \Theta \left( \frac{n}{2\alpha^3} \cdot m^{2\gamma-2} \cdot j^{-2\gamma} \right).
\]

(14)

To facilitate the application of the Maclaurin-Cauchy integral method, we further divided this case into three subcases, i.e., \( 0 < \gamma < 1/2 \), \( \gamma = 1/2 \), and \( 1/2 < \gamma < 1 \).

Subcase 1. \( 0 < \gamma < 1/2 \) and

\[
\rho_{sd} = \frac{1}{m} \Theta \left( \frac{n}{2\alpha^3} \cdot m^{2\gamma-2} \cdot j^{-2\gamma} \right)
\]

\[= \frac{1}{m^2} \Theta \left( \frac{n}{2\alpha^3} \cdot m^{2\gamma-1} \cdot \frac{1}{j^{2\gamma}} \right)
\]

\[= \Theta \left( \frac{n}{2\alpha^3} \cdot \frac{m^{2\gamma-1} - m^{2\gamma-2}}{1 - 2\gamma} \right).
\]

(15)

Subcase 2. \( \gamma = 1/2 \), \( \rho_{sd} = \Theta(n/(2m^3)) \cdot \log m/m \).

Subcase 3. \( 1/2 < \gamma < 1 \)

\[
\rho_{sd} = \Theta \left( \frac{n}{2m^3} \cdot \frac{2\gamma}{2\gamma - 1} \cdot m^{2\gamma-2} \right).
\]

(16)

Similarly, we can also determine the value when \( \gamma = 1 \) and \( \gamma > 1 \).

Case 2. \( \gamma = 1 \), \( p_j = \Theta(1/(j \log m)) \), and

\[
\rho_{sd} = \Theta \left( \frac{n}{2m^3} \cdot \frac{2}{\log^2 m} \right).
\]

(17)

Case 3. \( \gamma > 1 \), \( p_j = \Theta(j^{-\gamma}) \), and

\[
\rho_{sd} = \Theta \left( \frac{n}{2m^3} \cdot \frac{1}{2\gamma - 1} \right).
\]

(18)

The order meaning of the probability \( \rho_{sd} \) that the data packet can be directly delivered from the source node to the destination node is expressed as

\[
\rho_{sd} = \Theta \left( \frac{\beta}{2\gamma - 1} \right).
\]

(19)

where \( \beta = n/(ma^3) \).

We can get the capacity performance of 3D-MANET (with \( n \) nodes and \( m \) cells) under the two-hop nonredundant relay routing scheme by substituting equations (11) and (19) into the previously analyzed capacity expression equation (5), as

\[
C = n\lambda \leq \Theta \left( \frac{\beta m (1 + \log^2 m)}{2\gamma - 1} \right), \quad \gamma > 1,
\]

(20)
3.3. Two-Hop Nonredundant Relay Routing Scheme’s Network Delay. The processing delay in the end-to-end delay depends on the system’s hardware performance, and this part of the time is concise and can be ignored. Besides, a 3D space network with a specific height may ensure LOS transmission, and the influence of NLOS transmission can be approximately ignored, so the propagation time of data packets can also be ignored. Consequently, this paper primarily applies queuing theory to compute the queuing delay.

Since the nodes in the whole network obey the movement of i.i.d., the packet queue at the source node can be viewed as a Bernoulli/Bernoulli queuing model, in which the arrival rate of new packets each time slot is \( \lambda_j = \lambda \); a transmission opportunity occurs between nodes, which corresponds to a service opportunity in the Bernoulli/Bernoulli queue. At each time slot, the probability of such an opportunity occurring is the service rate at the source node, regardless of whether packets are waiting in the queue. Let \( \mu = \mu_1 + \mu_2 \), where \( \mu_1 \) represents the service rate at which the source node is scheduled to deliver data packets directly to the destination node; \( \mu_2 \) represents the service rate at which the source node is scheduled to send packets to the possible relay node.

According to the routing scheme described in Section 2.3, the average network delay can be divided into two parts, one is the average delay required to transmit data packets at the source node, denoted as \( E(\tau_s) \); another is the delay required for transmission through the path of “relay node → destination node,” denoted as \( E(\tau_r) \).

(1) Solve the average delay \( E(\tau_s) \) required to transmit packets at the source node

For a given time slot, the probability that the source node and the destination node meet in the cell \( j \) is \( p_j^2 \); there are \( np_j \) nodes in the cell \( j \), and each node will be randomly selected as the source node, that is, the probability that the source node can transmit data packets to the destination node is \( 1/(np_j) \); then, the service rate of the source node to the data packets to the destination node is \( \mu_1 = p_j^2 \cdot 1/(np_j) \).

Therefore, the average queue length of the data packets transmitted by the “source node → destination node” at the source node \( i \) is \( L_1 = (\rho_i (1 - \lambda_i))/((1 - \rho_i)) \), where \( \rho_i = \lambda_i/\mu_1 \). According to Little and Graves’ theorem [43], the average delay required for “source node → destination node” transmission is \( E(\tau_{sd}) = L_1 / \lambda_1 = (1 - \lambda_i)/(\mu_1 - \lambda_i) \). However, by comparing expressions (11) and (19), it is easy to simplify that the probability of the source node being scheduled for the “source node → destination node” transmission is very small, which indicates that the probability of the source node will be scheduled as the transmission mode of “source node → relay node → destination node” is a high probability; in this case, the main source of the average delay at the source node is mainly through the path “source node → relay node.” Correspondingly, the probability that the source node can transmit data packets to the relay node is \( 1/(np_j) \); then, the service rate of the source node to the data packets to the relay node is \( \mu_1 = 1/(np_j) \). Therefore, the average queue length of data packets used for the transmission of “source node → relay node” at source node \( i \) is \( L_2 = (\rho (1 - \lambda_i))/((1 - \rho)) \), where \( \rho = \lambda_i/\mu_2 \). According to Little and Graves’ theorem, the average delay required for “source node → relay node” transmission is \( E(\tau_{sr}) = E(\tau_{ur}) = L_2 / \lambda_1 = (1 - \lambda_i)/((\mu_2 - \lambda_i)) \).

(2) Solve the average delay \( E(\tau_r) \) required for the transmission of “relay node → destination node” at the relay node

For a given time slot, the output process of the source node \( i \) is also regarded as a Bernoulli process with rate \( \lambda_i \) [44], so the rate of data packets arriving from the source node to the relay node is \( \lambda_i = \lambda_i \cdot 1/(np_j) \). The probability that the relay node and the destination node meet in cell \( j \) is \( p_j^2 \), and the probability that the relay node can transmit data packets to the destination node is \( 1/(np_j) \); then, the service rate of the relay node to the data packet to the destination node is \( \mu_5 = p_j^2 \cdot 1/(np_j) \). Therefore, the packet queue at the relay node can be regarded as an M/M/1 queuing model with an arrival rate of \( \lambda_r \) and a service rate of \( \mu_5 \), and thus, the average queue length is \( L_3 = \lambda_i / ((\mu_5 - \lambda_r)) \) at the relay node. According to Little and Graves’ theorem, the average transmission delay required for the transmission of “relay node → destination node” is \( E(\tau_r) = L_3/\lambda_r = 1/(\mu_5 - \lambda_r) \).

Since the two path transmission modes of “source node → relay node” and “relay node → destination node” are triggered by equal probability, we can get the delay of 3D-MANET (with \( n \) nodes and \( m \) cells) under the two-hop nonredundant relay routing scheme by sorting, as

\[
E(\tau_{\text{sum}}) = \frac{1}{2} E(\tau_s) + \frac{1}{2} E(\tau_r)
\]

3.4. Two-Hop Redundant Relay Routing Scheme’s Network Capacity. Combined with the analysis idea of the network capacity of the nonredundant transmission scheme in Section 3.1, in this section, we directly deduce and give the capacity performance under the redundant transmission scheme. Assuming that the data packet still arrives at the node in the rate \( \lambda \), for a given time slot, each data packet will form \( r \) copies and forward it to \( r \) relay nodes, so each time slot network generates a new the total amount of packets is \( nr \lambda \). However, the total number of packets that the network can deliver is the same as in the nonredundant routing scheme, and hence, the following relationship holds:

\[
C_r = nr \lambda \leq m \bar{p}_{sd} + \frac{1}{2} C(\bar{p}_{\text{sum}} - \bar{p}_{sd}),
\]
i.e.,
\[ C_r = \frac{1}{r} \cdot \frac{1}{2} m(\overline{r}_p + \overline{r}_n). \]  

(23)

It can be observed that if the relay scheme with redundancy \( r \) is adopted, its network capacity will become \( 1/r \) under the nonredundant transmission scheme, i.e.,
\[ C_r = \frac{1}{r} C. \]  

(24)

3.5. Two-Hop Redundant Relay Routing Scheme’s Network Delay

Similar to the idea of analyzing the network delay of the nonredundant transmission scheme in Section 3.2, the average network delay is divided into two parts. Although the redundant relay routing strategy is adopted in this section, the source node and the destination node meet in the same cell, and completing the transmission’s probability does not change, so the time required for transmission through the path of the “source node \( \rightarrow \) destination node” continues to follow the analytical thinking of Section 3.2 and ignores it. This section will mainly analyze the delay \( \mathbb{E}(\tau_{rd}) \) required for transmission through the path of the “source node \( \rightarrow \) redundant relay node \( \rightarrow \) destination node.”

(1) Solve the average delay \( \mathbb{E}(\tau'_{rd}) \) required for the transmission of “source node \( \rightarrow \) relay node”

Similar to Section 3.2, the arrival rate of packets at source node \( i \) is \( \lambda_i \). The cell \( j \) contains \( np_j \) nodes, and each node will be randomly selected as the source node; accordingly, the probability that the source node can transmit data packets to the first relay node is \( r/(np_j) \), and then, the service rate of the source node to the first relay node is \( \mu_1 = r/(np_j) \).

Next, the source node’s service rate for packets destined for the remaining \( (r - 1) \) relay nodes becomes \( \mu'_1 = 1/(r - 1) \), which means that once the node is selected as the source node, it will send all redundant packets. Therefore, the packet queue at source node \( i \) can be viewed as a Bernoulli/Bernoulli queueing model with arrival rate \( \lambda_i \) and service rate \( \mu'_1 \) and a Bernoulli/Bernoulli queueing model with arrival rate \( \lambda_i \) and service rate \( \mu'_1 \). Thus, the average length of the queue at the source node \( i \) is \( L'_1 + L''_1 = (\rho'/(1 - \lambda_i)) + (\rho''/(1 - \lambda_i)) \), where \( \rho' = \lambda_i/\mu'_1 \) and \( \rho'' = \lambda_i/\mu'_2 \). According to Little and Graves’ theorem, the average delay required for “source node \( \rightarrow \) relay node” transmission is \( \mathbb{E}(\tau'_{rd}) = (L'_1 + L''_1)/\lambda_i = (1 - \lambda_i)/(\mu'_1 - \lambda_i) + (1 - \lambda_i)/(\mu'_2 - \lambda_i). \)

(2) Solve the average delay \( \mathbb{E}(\tau_{rd}) \) required for the transmission of “redundant relay node \( \rightarrow \) destination node”

Similar to Section 3.2, the output process of the source node \( i \) is also regarded as a Bernoulli process with a rate of \( \lambda_i \), so the rate of data packets from the source node to the relay node is \( \lambda_r = \lambda_i \cdot 1/(np_j) \). Due to the existence of \( r \) relay nodes, the probability of the relay node meeting the destination node in cell \( j \) increases, from \( p'_j \) to \( p''_j \), the probability that the relay node can transmit data packets to the destination node is \( 1/(np_j) \), and then, the service rate of the relay node to the data packets to the destination node is \( \mu'_2 = p''_j \cdot r/(np_j) \). Therefore, the packet queue at the relay node can be regarded as an M/M/1 queuing model with an arrival rate of \( \lambda_r \) and a service rate of \( \mu'_2 \). The average queue length at the relay node is \( L_3 = \lambda_r/\mu'_2 = 1/(\mu'_2 - \lambda_r) \). According to Little and Graves’ theorem, the average transmission delay required for “relay node \( \rightarrow \) destination node” transmission is \( \mathbb{E}(\tau_{rd}) = L_3/\lambda_r = 1/(\mu'_2 - \lambda_r) \).

After sorting, the average network delay results can be obtained as follows:

\[ \mathbb{E}(\tau'_{sum}) = \mathbb{E}(\tau'_{rd}) = \mathbb{E}(\tau'_n) + \mathbb{E}(\tau'_{rd}) = \frac{1}{2} \mathbb{E}(\tau'_{rd}) + \mathbb{E}(\tau'_n) = \frac{1}{2} \frac{1 - \lambda_i}{\mu'_1 - \lambda_i} + \frac{1 - \lambda_i}{\mu'_2 - \lambda_i} + \frac{1}{\mu'_2 - \lambda_i} \]

\[ = \frac{1}{2} \left[ n \cdot \frac{p'_i - p'_j \lambda_i}{1 + (np_j \lambda_i)/r} + (r - 1) \frac{1 - \lambda_i}{1 - \lambda_i / (r - 1)} + \frac{n}{r} \cdot \frac{p'_i}{p''_j - \lambda_i} \right]. \]  

(25)

Note that when the packet redundancy \( r \) is small (or \( r = o(n) \)), formula (25) can be further written as

\[ \mathbb{E}(\tau'_{sum}) = \frac{1}{2} \Theta(n + \frac{n}{r}) = \Theta(n^2 + \frac{n}{r}). \]  

(26)

However, when \( r \) is large (or \( r \rightarrow \Theta(n) \)), its own influence on network delay cannot be ignored; formula (25) can be further written for

\[ \mathbb{E}(\tau'_{sum}) = \frac{1}{2} \Theta(r + \frac{n}{r}). \]  

(27)

To minimize the average delay, use the average value inequality to optimize formula (27) further:

\[ \min_{r \geq 1} \mathbb{E}(\tau'_{sum}) = \min_{r \geq 1} (r + \frac{n}{r}). \]  

(28)

Obviously there is \( r + (n/r) \geq 2 \sqrt{n} \), if and only if \( r = \sqrt{n} \), the equal sign is established, and the optimal network delay \( \mathbb{E}(\tau'_{sum}) = \sqrt{n} \) is achieved at this time; correspondingly, the optimal packet redundancy \( r \) should be set to \( r = \sqrt{n} \) (or \( r = \Theta(\sqrt{n}) \)).
4. Capacity and Delay Performance Optimization of 3D-MANET

Through the analysis in Section 3, it can be seen that the delay can be reduced to $\frac{1}{\sqrt{n}}$ under the nonredundant transmission scheme through the data packet redundancy technology. However, while this effectively optimizes the time delay performance of 3D-MANET, its capacity is reduced to $\frac{1}{r}$, which is unacceptable. In addition, from the theoretical results shown by formula (20), it can be seen that the delay performance of this paper is slightly reduced compared with the study obtained by using the node uniformly distributed mobility model. One of the main reasons is that network cells with high node density have tremendous pressure to transmit data. Second, from the reverse perspective, it is precisely because these dense cells contain more nodes, and the classic two-hop nonredundant relay routing strategy essentially still enables nodes to transmit data in a peer-to-peer form so that more nodes in the cell are in an idle state, thereby lowering the overall transmission efficiency of this part of the network cell.

4.1. The WAP-Enabled Multihop Relay Routing Scheme. This paper alleviates transmission pressure by deploying wireless access points (WAPs) in cells with a high degree of node aggregation (hereinafter referred to as “dense cells”), taking into account these negative factors. On this basis, this paper proposes a WAP-enabled routing scheme, which allows any node in a dense cell to access WAP in one hop and can use WAP to broadcast data packets in a polling manner, consequently accelerating the delivery process of data packets, reducing network delay, and improving network performance. Before designing this routing scheme, the first problem we need to solve is to determine whether a cell is a dense cell. If the number of nodes in a cell is not less than $n_T$, it is referred to as a dense cell. Further assuming that the number of dense cells in the entire network is $E(m_T)$. The following relationship exists between $n_T$ and $E(m_T)$ according to formula (1).

$$n_T = n \cdot F(m) \cdot E^{-\gamma}(m_T).$$ (29)

The overall idea of this routing strategy using WAP is shown in Figure 6. The data packets have the following possible paths from the source node to the destination node:

1. If the source node and the destination node are in the same cell, the data packet can be directly transmitted by the “source node $\rightarrow$ destination node” path.

2. When the source and the destination nodes are not in the same cell, then first determine whether the unit is a dense unit
   a. If it is a dense unit, the source node will connect one hop to the WAP and broadcast data packets to the relay nodes (These relay nodes are referred to as I-type relay nodes and call them as II-type relay nodes of nondense cells’ relay nodes. In different time slots, all nodes may become one of the two types of relay nodes. Here, they are temporarily classified for the convenience of description and analysis. All I-type relay nodes are possible to carry a copy of the packet.) by polling. That is, the data packet is delivered by the path of “source node $\rightarrow$ WAP $\rightarrow$ I-type relay node $\rightarrow$ destination node”.
   b. If it is a nondense cell, the source node transmits the data packet to the II-type relay node. Here, similar to the data packet transmission process of the two-hop nonredundant relay routing scheme. Then, as the network node moves, when the II-type relay node moves to the dense cell, it can broadcast the data packet in the cell by accessing the WAP and further transmit the data packet through the redundant I-type relay node that receives the data packet in the cell to the destination node. In addition, the II-type relay node can also be directly delivered to the
destination node, provided that they are in the same cell. The analysis in Section 3.2 demonstrates that the probability of the sending node and the receiving node moving to the same cell simultaneously is small. According to Zipf’s law, the dense cells in the network are much less than the nondense cells, so most of the data packets in the whole network will follow the path “source node → II-type relay node (nondense cell) → WAP → I-type relay node (dense cell) → destination node” completes the transmission.

Essentially, this WAP-enabled multihop relay routing scheme combines the advantages of broadcast transmission and packet redundancy techniques.

4.2. The WAP Relay Routing Scheme’s Network Capacity.

When analyzing the delay performance of the two-hop nonredundant relay routing scheme in Section 3.2, we pointed out that the packet queue at the source node can be regarded as a Bernoulli/Bernoulli queuing model, in which the arrival rate of the packet queue at the source node can be regarded as a Bernoulli/Bernoulli queuing model. In Section 3.2, we pointed out that the path “source node → II-type relay node (nondense cell) → WAP → I-type relay node (dense cell) → destination node” completes the transmission.

According to the transmission scheme shown in Figure 6, there are three possible transmission schemes at the source node: “source node → destination node,” “source node → II-type relay node,” and “source node → WAP.” Let $v_1$ denote the transmission rate at which the source node is scheduled to deliver data packets directly to the destination node, and $v_2$ denotes the transmission rate at which the source node is scheduled to transmit data packets to II-type relay nodes or access WAP. Since relay nodes or WAPs are scheduled to perform input and output operations with equal probability, the output rate of broadcasting in dense cells by accessing WAP is $v_2$, and the output rate of the I-type relay node and the II-type relay node to deliver the data packet to the destination node is $v_2$. Therefore, the total rate of transmission opportunities for the entire network is $v = v_1 + 3v_2$. According to the transmission mode described in Section 2.2, a transmission opportunity exists for a given time slot when there are at least two nodes in the same cell; that is, a packet delivery event occurs. If the cell $j$ has at least two nodes and the probability of one transmission generation is $p^j_{\text{sum}}$, then the following relationship holds:

$$\sum_{j=1}^{m} p^j_{\text{sum}} = n(v_1 + 3v_2).$$

(30)

Similarly, when there is at least one source-destination node pair in the same cell, it is considered that there is a direct transmission opportunity between the source node and the destination node. And the probability that the number of source-destination node pairs in cell $j$ satisfies $E(\text{SD}) \geq 1$ is $p^j_{\text{id}}$; then, the following relationship holds:

$$\sum_{j=1}^{m} p^j_{\text{id}} = nv_1.$$  

(31)

Simultaneous formula (30) and formula (31), we can get

$$v_1 = \frac{mp_{\text{id}}}{n},$$  

(32)

$$v_2 = \frac{m(p_{\text{sum}} + p_{\text{id}})}{3n}.$$  

Therefore, the network capacity under the premise of ensuring stable network transmission can be obtained as

$$C_w = n\lambda \leq n\nu = n(v_1 + 3v_2) = \frac{1}{3} m(p_{\text{sum}} + 2p_{\text{id}}).$$

(33)

We can get the capacity performance of 3D-MANET (with $n$ nodes and $m$ cells) under the WAP-enabled relay routing scheme by substituting equations (11) and (19) into the capacity expression equation (33), as

$$C_w = n\lambda \leq \begin{cases} \Theta \left( \frac{m - m^\gamma}{3 - 3\gamma} + \frac{1 - m^{\gamma + 1}}{3 - 3\gamma} \right), & 0 < \gamma < \frac{1}{2}, \\ \Theta \left( \frac{m - m^\gamma}{3 - 3\gamma} - \frac{\log m}{3} \right), & \gamma = \frac{1}{2}, \\ \Theta \left( \frac{m^{\gamma + 1} - m^\gamma}{3 - 3\gamma} + \frac{2y}{6y - 3} \cdot m^{\gamma + 1} \right), & \frac{1}{2} < \gamma < 1, \\ \Theta \left( \frac{m^{\gamma + 1} - m^\gamma}{3 - 3\gamma} + \frac{3m^\gamma}{3\log m} \cdot \log m \right), & \gamma = 1, \\ \Theta \left( \frac{3\gamma - 2}{(2\gamma - 1)(\gamma - 1)} \cdot m^{\gamma + 1} \right), & \gamma > 1, \end{cases}$$

where $\beta = n/(ma^3)$.

4.3. The WAP Relay Routing Scheme’s Network Delay. For a given packet, it has the highest probability of being delivered by the path of “source node → II-type relay node → WAP broadcast → I-type relay node → destination node,” in which the transmission path determines the average delay $E(r^\infty_{\text{sum}})$ of the network, which mainly includes the following three parts (similar to the two-hop routing scheme under the network delay performance analysis, the repeated parts will not be described in detail).

1. The average delay required for the source node to transmit data packets to the II-type relay node is $E(r^\infty_{\text{sr}}) = \frac{L_4}{(1 - \lambda_i)}(1 - (1/(n\lambda_i)) - \lambda_i)$; since the values are the same, the result of Section 3.2 is used, where $L_4$ represents the average queue length at the source node for “source node → II-type relay node” transmission.

2. Solve the average delay $E(r^\infty_{\text{rig}})$ required for “II-type relay nodes broadcast data packets to I-type relay nodes through WAP.”

In a nondense cell, the probability that the source node is scheduled to transmit to an II-type relay node is $v_2/n$, and the
remaining \((n - 2)\) nodes in the network may become II-type nodes relay node, so that the rate at which packets arrive at the II-type relay node is \(\lambda_{n} = \lambda_{1} \cdot v_{2} \cdot 1/(n-2)\). Next, the II-type relay node takes the probability \(\mu_{n} = (n_{T}E(m_{T})/n \cdot v_{2}/n_{T})\) broadcasts packets in dense cells, where \(n_{T}E(m_{T})/n\) indicates the probability that the II-type relay node moves to a dense cell; \(v_{2}/n_{T}\) represents the probability that a II-type relay node obtains a chance to broadcast a packet through WAP. In addition, since the arrival and output processes of data packets in the II-type relay node are independent of each other, the discrete-time Markov chain formed by the queue can be regarded as a birth-death chain, which is related to the input rate of \(\lambda_{n}\) and service rate \(\mu_{n}\). It is the same as the continuous time M/M/1 queue. Thus, the average queue length at the II-type relay node is \(L_{n} = \lambda_{n} / (\mu_{n} - \lambda_{n})\).

According to Little and Graves' theorem, the average delay required for “II-type relay node \(\rightarrow\) I-type relay node” transmission is \(E(t_{sum}^{w}) = L_{n} / \lambda_{n} = 1/(\mu_{n} - \lambda_{n})\).

(3) Solve the average delay \(E(t_{sum}^{w})\) required for “I-type relay node transmits data packets to destination node”

The probability that a II-type relay node transmits to a I-type relay node by broadcasting is \(n_{T}/(n-2)\), so the probability of a data packet reaching a I-type relay node is \(\lambda_{1} = \lambda_{1} \cdot n_{T}/(n-2)\). Next, the I-type relay node starts with \(\mu_{n} = n_{T} \cdot v_{2} \cdot \sum_{i=1}^{m} \left(p_{i} \cdot 1/(n \cdot p_{i})\right)\) to deliver the data packet to the destination node, where \(n_{T} \cdot v_{2} \cdot \sum_{i=1}^{m} \left(p_{i} \cdot 1/(n \cdot p_{i})\right)\) means the probability that a I-type relay node in the network can deliver a packet to the destination node. Similarly, the packet queue at a I-type relay node can be regarded as an M/M/1 queuing model with an arrival rate of \(\lambda_{n}\) and a service rate of \(\mu_{n}\), so that the average queue length at a I-type relay node is \(L_{n} = \lambda_{n} / (\mu_{n} - \lambda_{n})\). According to Little and Graves' theorem, the average delay required for the transmission of “II-type relay node \(\rightarrow\) I-type relay node” is \(E(t_{sum}^{w}) = L_{n} / \lambda_{n} = 1/(\mu_{n} - \lambda_{n})\).

Substituting and arranging the above results, we get

\[
E(t_{sum}^{w}) = \Theta \left( \frac{n}{E(m_{T})} \right) + \Theta \left( \frac{n}{n_{T}} \right) > \Theta \left( \frac{2}{E(m_{T})} \cdot \frac{n}{n_{T}} \right) \]

\[
\geq \Theta \left( \frac{2n^{2}(1+y)}{E(m_{T}) \cdot (1+y)} \right). \tag{36}
\]

If and only if \(E(m_{T}) = n_{T} = \Theta(nF(m))^{1/(1+y)}\), the equal sign is established, and the delay gets the minimum value.

We can get the delay of 3D-MANET (with \(n\) nodes and \(m\) cells) under the two-hop nonredundant relay routing scheme by sorting, as

\[
E(t_{sum}^{w}) \begin{cases} \Theta \left( \frac{2n^{2}(1+y)}{E(m_{T}) \cdot (1+y)} \right), & y > 1, \\ \Theta \left( \frac{2n^{2}(1+y)}{E(m_{T}) \cdot (1+y)} \right), & y = 1, \\ \Theta \left( \frac{2n^{2}(1+y)}{E(m_{T}) \cdot (1+y)} \right), & y < 1, \end{cases}
\]

5. Numerical Approximation Results and Discussion

This section mainly conducts a comparative analysis of these performance results and discusses the impact of different parameters on the network performance in order to draw systematic conclusions. Before this, the influence of node aggregation degree distribution index \(y\) on node movement distribution is first shown. As shown in Figure 7, it can be seen that when \(y\) is smaller, the distribution of nodes in the entire network is more uniform. As \(y\) increases, nodes tend to cluster in a few "hot" areas, while nodes in most network units show a sparse distribution.

5.1. Influence of Parameters \(y\) on Network Capacity Performance. It can be seen from formula (24) that if the two-hop relay scheme with redundancy \(r\) is adopted, its capability will become \(1/r\) of the nonredundant relay transmission scheme, and this performance is unacceptable. Given this, this subsection focuses on comparing the capacity performance under the two-hop nonredundant relay and the proposed WAP relay scheme, referring to the theoretical results described in formulas (20) and (34), to intuitively analyze the influence of node aggregation degree distribution index \(y\) on network capacity; combined with the actual situation, consider the large-scale network and set the number of network nodes \(n = 15000\) and the number of network cells \(m = 216 = 6^{3}\); the value range of \(a\) can determine \(a = 2\), and can also set these parameters to other reasonable values, and then get progressive numerical results, as shown in Figure 8.

The network capacity of using the WAP relay is slightly lower than that of the two-hop nonredundant relay strategy. With the change of the value of \(y\), the trend of capacity under the two schemes is comparable. When \(0 < y < 1\), the capacity increases approximately exponentially with the increase of \(y\). \(y = 1\) is the discontinuity point of the capacity function \(C(n, m, a, y)\). From the
point of view of the function order and practical significance, it can be considered that the network capacity is reached the maximum when $\gamma \approx 1$. When $\gamma > 1$, the capacity delay rapidly with the increase of $\gamma$. Referring to Figure 7, the more cells in the network that can transmit at the same time, the greater the network capacity. If $\gamma$ continues to rise, the nodes tend to be clustered in dense cells, and the nodes will be more sparsely distributed in nondense cells; then, there will be an increase in the number of free cells. Concurrently, the network traffic load of dense cells is excessively high, resulting in a steady decline in network capacity.

**Figure 7:** Illustration of the distribution of nodes when $\gamma$ takes different values.
5.2. Influence of Parameters \( r, n, \) and \( \gamma \) on Network Delay Performance. From the network delay performance derived and analyzed in Sections 3.2 and 3.5, the network delay performance of the two-hop relay transmission scheme is dependent on the network scale and the degree of data packet redundancy. If the nonredundant relay scheme is adopted, that is, \( r = 0 \), as shown in Figure 9, the number of nodes in the network is set to be at most 10000. With the increase of the network scale (i.e., the number of network nodes \( n \)), the delay increases linearly to \( \Theta(n) \). If a redundant scheme is used, and \( r = o(n) \), as shown in Figure 9, with the gradual increase of \( r \), the delay can be reduced to \( \Theta(n/2) \). If the redundant scheme is adopted, and \( r \to \Theta(n) \), as shown in Figure 10, the best delay performance is attained when \( r = 100 \), and increasing \( r \) will not improve the delay, and at this time, the network capacity will drop sharply with the increase of \( r \) value.

It can be observed that the delay under this routing strategy is significantly affected by \( \gamma \) referring to the delay results of the WAP relay scheme described in formula (37), and also set the number of network nodes \( n = 15000 \) and the number of network cells \( m = 216 = 6^3 \) and \( \alpha = 2 \); the numerical results shown in Figure 11 are obtained. Specifically, when \( 0 < \gamma < 1 \), the delay decreases with the increase of \( \gamma \). From the point of view of function order and practical significance, the delay is the smallest when \( \gamma = 1 \); when \( \gamma > 1 \), the delay increases with the increase of \( \gamma \). From the analysis in Section 4.3, the network delay expression of the WAP relay is

\[
E(t^n) = \Theta \left( \frac{n}{E(m_T)} \right) + \Theta \left( \frac{n}{n_T} \right),
\]

where \( \Theta(\frac{n}{E(m_T)}) \) represents the delay required by the II-type relay node to broadcast data packets to the I-type relay node through WAP and \( \Theta(\frac{n}{n_T}) \) represents the delay required by the I-type relay node to transmit the data packet to the destination node. When \( 0 < \gamma < 1 \), with the increase of \( \gamma \), the network nodes tend to move to the dense cell. Correspondingly, the number of nodes \( n_T \) in the dense cell increases, and the number of dense
cells \( E(m_T) \) in the network is also increasing, so the network delay is decreasing. When \( \gamma > 1 \), with the increase of \( \gamma \), most nodes tend to gather in a few dense cells; correspondingly, \( E(m_T) \) becomes smaller as delay increases. If and only if \( E(m_T) = n_T \), then \( \gamma = 1 \), \( E(m_T) \) and \( n_T \) reaches the maximum value, so the network delay is the smallest.

The delay performance of WAP multihop relay routing is much superior to that of two-hop nonredundant relay routing under the premise of less network capacity loss performance. When \( \gamma \in (0, 6) \), the delay performance is also better than the two-hop redundant (\( r = o(n) \)) relay scheme. The scheme takes full use of the node resources of dense cells, significantly improves the transmission efficiency, and
5.4. Influence of Parameter \( m \) on Network Delay/Capacity Performance. As mentioned above, the WAP multihop relay routing algorithm can achieve better network performance. When \( \gamma = 1 \), the delay and capacity performance is the best, and the delay and capacity are also satisfied with the best trade-off relationship. Based on this, this section continues to discuss the impact of the number of network cells \( m \) on “delay/capacity” \( \frac{E(t_{\text{sum}})}{C_w} \) performance under the WAP relay routing scheme. Also, set \( n = 15000 \) and \( \alpha = 2 \), and get the numerical results shown in Figure 12. It is clear that when the number of network cells is in \( m \in [210,300] \), the optimal network performance can be further achieved on the basis of the best trade-off between network delay and network capacity achieved at \( \gamma = 1 \). It should be noted that increasing the number of network cells \( m \) beyond this range has no significant gain on network performance. When \( m \) is within this range, it is reasonable and in line with the actual situation.

5.3. Influence of Parameter \( m \) on Network Delay/Capacity Performance. As mentioned above, the WAP multihop relay routing algorithm can achieve better network performance. When \( \gamma = 1 \), the delay and capacity performance is the best, and the delay and capacity are also satisfied with the best trade-off relationship. Based on this, this section continues to discuss the impact of the number of network cells \( m \) on “delay/capacity” \( \frac{E(t_{\text{sum}})}{C_w} \) performance under the WAP relay routing scheme. Also, set \( n = 15000 \) and \( \alpha = 2 \), and get the numerical results shown in Figure 12. It is clear that when the number of network cells is in \( m \in [210,300] \), the optimal network performance can be further achieved on the basis of the best trade-off between network delay and network capacity achieved at \( \gamma = 1 \). It should be noted that increasing the number of network cells \( m \) beyond this range has no significant gain on network performance. When \( m \) is within this range, it is reasonable and in line with the actual situation.

5.4. Influence of the Parameter \( n \) on the Single Node Capacity and Single Node Delay. As shown in Figure 13, under the WAP relay routing strategy, the increase in the number of network nodes \( n \) intensifies the competition for transmission resources between nodes, and the capacity of a single node continues to decline. When the network scale tends to infinity, the capacity of a single node approaches zero. At the same time, as the number of nodes increases, the probability of encounters between the source node, I-type relay node, II-type relay node, and destination node continues to increase; that is, the transmission opportunities between nodes increase, and the single node delay continues to decrease.

6. Conclusions and Further Work

In this paper, we investigate the scaling law of 3D-MANET capacity and delay under different routing schemes. This study is compatible with the earlier theoretical results in 2D and 3D scenarios, and the network performance is related to parameters such as network scale, packet delivery rate, and redundancy. Meanwhile, it is also an extension of the previous, and our results can cover a variety of distribution scenarios with the change of \( \gamma \) value.

Since this paper considers a large-scale multihop 3D-MANET network composed of any communication entities on land, sea, or air, these general theoretical results can be applied to practical scenarios: dense viaducts, overpasses, and highway scenarios on the Internet of vehicles; air-space-ground integrated self-organizing network scenarios including aircraft, UAV, ground base stations, fleet, and other communication entities; UAV swarm network scenarios; etc.

In the future, one valuable direction is to further explore the 3D-MANET performance under the multicast routing strategy for obtaining a more systematic and comprehensive conclusion. In addition, the development of machine learning and artificial intelligence technology is in full swing [45, 46]. This paper has shown that a reasonable division of network cells can help improve network performance. By designing experiments, we can explore the optimal machine learning algorithm on data sets to cluster network nodes reasonably to indirectly change nodes’ spatial aggregation and distribution to achieve optimal network performance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported in part by the Wireless Ad Hoc Network Communication Performance Testing Research Project under Grant W21GY1300010 and in part by the Key Program of the National Natural Science Foundation of China under Grant 61931001.
References


