



# Article A Multipath Data-Scheduling Strategy Based on Path Correlation for Information-Centric Networking

Yong Xu<sup>1,2</sup>, Hong Ni<sup>1,2</sup> and Xiaoyong Zhu<sup>1,2,\*</sup>

- <sup>1</sup> National Network New Media Engineering Research Center, Institute of Acoustics, Chinese Academy of Sciences, No. 21, North Fourth Ring Road, Haidian District, Beijing 100190, China; xuy@dsp.ac.cn (Y.X.); nih@dsp.ac.cn (H.N.)
- <sup>2</sup> School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, No. 19(A), Yuquan Road, Shijingshan District, Beijing 100049, China
- \* Correspondence: zhuxy@dsp.ac.cn; Tel.: +86-131-2116-8320

Abstract: Information-Centric Networking (ICN) has revolutionized the manner of content acquisition by shifting the communication mode from host-centric to information-centric. Considering the existing, large amount of IP infrastructure in current networks, the new ICN architecture is proposed to be compatible with existing networks in order to reduce deployment cost. However, due to compatibility with IP networks, ICN data packets must be transmitted through the default path provided by IP routing regulations, which also limits the transmission efficiency and reliability of ICN. In order to address this issue, this paper introduces a multipath transmission method applied in ICN which takes full advantage of the functions and characteristics of ICN and builds multiple end-to-end relay paths by using the ICN routers as relay nodes. We then propose a relay-node-selection algorithm based on path correlation to minimize the impact of overlapping links. Moreover, we comprehensively calculate the path state value by combining the round-trip time and packet loss rate and propose a multipath data-scheduling algorithm based on the path state value. Simulation experiments show that the proposed method can maintain high bandwidth utilization while reducing the number of out-of-order packets.

Keywords: ICN; multipath transmission; path correlation; data scheduling

# 1. Introduction

Since its birth, the Internet has become one of the most important ways for users to obtain information resources. In the past decade, a series of highly significant innovations and reforms have taken place in communication technology and network infrastructure. With the emergence of new network applications and services such as the industrial Internet, big data, artificial intelligence, blockchain, virtual reality, etc., Internet users and traffic have also experienced explosive growth. Information sharing has become the main function and goal of the Internet, and the usage mode of the Internet has gradually evolved toward content access and information acquisition. Users no longer pay attention to the location of the content but to the content itself, as well as the speed, quality, and security of content retrieval and transmission. The traditional host-oriented communication mode is not able to effectively meet the current content-centric requirements and is unable to provide efficient content distribution services. In order to solve this problem, researchers propose to establish overlay networks through patching at the application layer to enhance content distribution capabilities such as Content Delivery Networks (CDN) and Peer-to-Peer (P2P) [1]. Although this patch-based method alleviates the pressure of content demand expansion to some extent, its superimposed improvement leads to problems of low network resource utilization and high complexity.

In this context, the Information-Centric Networking (ICN) is proposed, which shifts the communication mode from host-centric to information-centric, completely changing the



Citation: Xu, Y; Ni, H.; Zhu, X. A Multipath Data-Scheduling Strategy Based on Path Correlation for Information-Centric Networking. *Future Internet* 2023, *15*, 148. https://doi.org/10.3390/fi15040148

Academic Editor: Cheng-Chi Lee

Received: 13 March 2023 Revised: 3 April 2023 Accepted: 5 April 2023 Published: 11 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). retrieval and transmission of content [2]. Since its proposal, ICN has attracted the attention of many research institutions and individuals. Different from existing TCP/IP networks, the content identifiers (ID) and locators are separated in ICN. According to the difference between addressing and routing methods, the existing ICN architecture can be divided into two types. The first is the ICN paradigm of name-based routing, such as Content Centric Networking (CCN) [3], Named Data Networking (NDN) [4,5], and Publish–Subscribe Internet Technology (PURSUIT) [6], which use hierarchical and aggregable names, and their addressing and routing are coupled. However, this type of ICN paradigm cannot take advantage of existing IP infrastructure, resulting in high deployment costs. The latter is the ICN paradigm of stand-alone name resolution, such as MobilityFirst [7], Network of Information (NetInf) [8] and on-Site, Elastic, Autonomous Network (SEANet) [9], which use flat and non-semantic names and are decoupled in terms of addressing and routing. This type of ICN paradigm is compatible with existing IP networks, thus greatly reducing deployment costs. The multipath transmission scheduling strategy proposed in this paper is also applied in the latter ICN paradigm.

The ICN paradigm of stand-alone name resolution uses ID-based naming and addressing, enhancing the addressing space and diversity. However, due to compatibility with IP networks, the ICN packets are still transmitted along the default path provided by the IP routing protocol. With the increase in network traffic, the congestion probability of the default path will increase, and the end-to-end throughput will decrease. This mismatch between addressing and routing limits the transmission quality of ICN. In order to improve the transmission efficiency and robustness, researchers have proposed various multipath transmission methods. At the network layer, Equal-Cost Multi-Path (ECMP) [10] is a classic multipath routing protocol which is widely used in Data Center Networks (DCNs). However, the equal traffic distribution makes ECMP difficult to adapt to asymmetric links. Segment routing [11] encodes the transmission path in the packet header, making the transmission path definable. Particularly when it is used with software-defined networking (SDN), the performance of segment routing has been effectively improved [12]. However, the excessive overhead limits the transmission efficiency. At the transport layer, most multipath transmission methods establish multiple parallel paths through multiple network interfaces of the terminal device to improve network bandwidth utilization [13,14]. However, these methods can only be applied in the scenario in which the terminal device has a multi-homing feature. At the application layer, researchers propose deploying overlay networks to construct virtual paths to enhance the flexibility and efficiency of end-to-end transmission [15,16]. However, these methods suffer from low network resource utilization and high complexity due to the chimney-style overlay deployment. In addition, multipath transmission technology has been widely studied in Wireless Sensor Networks (WSN) and Network on Chip (NoC) networks. In WSN, researchers have proposed some multi-hop relay routing algorithms to improve network performance and reduce energy consumption [17,18]. In order to improve the fault tolerance and recovery ability of NoC, researchers propose to apply the self-organization algorithm in WSN to NoC and provide the corresponding rules to bypass the blocked network sections [19]. Although such algorithms can enhance transmission performance, we believe that mechanically applying them to ICN cannot achieve significant performance gains.

In this context, we introduce a multipath transmission method combining the functions and characteristics of ICN networks and design a scheduling strategy based on path correlation to enhance transmission efficiency and robustness. The main contributions of this paper are as follows:

- We introduce a multipath transmission method applied to ICN which establishes multiple end-to-end relay paths by treating ICN routers as relays between the source and the destination.
- We propose a relay-node-selection algorithm based on path correlation to minimize the impact of overlapping links. Moreover, we comprehensively calculate the path

state value by combining round-trip time (RTT) and packet loss rate and propose a multipath data-scheduling algorithm based on the path state value.

• We evaluate the proposed method through a series of simulation experiments. Experimental results show that the proposed method can maintain high bandwidth utilization while reducing the number of out-of-order packets.

The rest of this paper is organized as follows. In Section 2, we discuss the related technologies of multipath transmission technologies and data-scheduling strategies. Then, the ID-based multipath transmission method is introduced in Section 3. Section 4 presents a relay-node-selection algorithm and multipath data-scheduling algorithm. Simulation experiments are carried out in Section 5. In Section 6, we finally conclude our contributions.

# 2. Related Works

Multipath transmission technologies have been extensively studied due to their ability to improve transmission robustness and efficiency. In this section, we will provide a detailed introduction to the multipath transmission technologies and multipath data-scheduling strategies.

#### 2.1. Multipath Transmission Technology

ECMP [10] has been widely used in DCNs due to its simplicity and low deployment cost. It distributes traffic to equivalent multiple paths, achieving multipath load balancing. However, this equal distribution of traffic makes its performance poor in asymmetric network topologies. Weighted Cost Multi-Path (WCMP) [20] can distribute traffic proportionally among multiple links, but its scalability is poor and can only be deployed in medium or small networks. In addition, segment routing [11] encodes the transmission path in the data packet header, making the transmission path definable. Especially after the introduction of SDN technology, the performance of segment routing has been effectively improved. In contrast to the above methods, some researchers have proposed the concept of overlay routing. The overlay network builds virtual paths on the underlying physical network and defines virtual topology and routing patterns based on application requirements. Overlay routing focuses on how to establish routing rules among the virtual nodes in an overlay network to customize the extension of Internet services. Detour [15] deploys overlay nodes in the network called smart gateways, which probe the delay, packet loss rate, and other metrics on the underlying physical network in real time and then forward the data through tunnels to bypass the faulty path. In [16], the authors propose the Resilient Overlay Network (RON), which enables the quick detection and recovery of fault paths. RON employs a full network connection and periodically probes the status of virtual links between overlay nodes. The network topology and path information are distributed to each neighboring node through a link-state routing protocol similar to Open Shortest Path First (OSPF) [21]. Once the current path is congested or fails, it can quickly switch to the optimal alternative relay path. Through experiments, RON found that using only singlehop indirect routing through a single intermediate node can yield most of the benefits in terms of latency and packet loss. Therefore, selecting appropriate relay nodes between the source and destination to exploit path diversity to bypass faults or congestion has drawn the attention of researchers. Han et al. [22] presented a clustering-based method which clusters the overlay nodes for which the overlap rate of relay paths between the source and destination exceeds a threshold. Randomly selecting only one overlay node in the same cluster increases the difference between relay paths, thereby enhancing the ability to avoid failures. Similarly, the authors in [23] proposed an earliest-diversity heuristic algorithm which selects the relay node that is separated from the current path earliest as the next hop. It is based on the assumption that the earlier a path separates from the original path, the less likely it is to intersect with the original path again. In addition, Roy et al. [24] used a heuristic algorithm to increase the overlay nodes incrementally so that the link overlap rate between relay paths was minimal. In [25], the author proposed to preferentially select overlay nodes with less failure probability to build relay paths. This method assumes

that the failure probabilities of all links in the network are known conditions, which is obviously unrealistic.

### 2.2. Multipath Data-Scheduling Strategy

Multipath data scheduling refers to distributing traffic to different network paths according to specific strategies during transmission to achieve load balancing. The roundrobin scheduling strategy distributes the application layer's data to each sub-path in a polling manner, ignoring the path state differences such as bandwidth and RTT. This strategy performs well when the state differences among the sub-paths are not significant; however, when the state differences are significant, it may lead to a decrease in the overall throughput. In contrast, the weighted round-robin scheduling strategy takes into account the status differences between paths. The Lowest RTT First algorithm [26] is a typical weighted round-robin scheduling strategy that prioritizes multiple available paths according to RTT. The smaller the RTT of the path, the higher its priority. For instance, Multipath TCP (MPTCP) [13] integrates the Lowest RTT First algorithm in the Linux kernel. Considering that the transmission paths of data packets and acknowledge packets are usually asymmetric, RTT cannot accurately reflect the real path state. Therefore, Forward-Delay-Based Packet Scheduling (FDPS) [27] prioritizes the path by measuring the forward delay of the data packets, thus improving the accuracy of the path state estimation. The redundant scheduling strategy [28] distributes packets repeatedly to each sub-path, alleviating the out-of-order problem to a certain extent. However, the same packets are transmitted over different paths, which affects the resource utilization of multiple paths. The hash scheduling strategy takes information such as the source IP address, destination IP address, source port and destination port as inputs to a hash function and selects the path based on the output result. For instance, ECMP applies a hash scheduling strategy [29]. However, the hash scheduling strategy is prone to hash collision, making it difficult to achieve load balancing. The prediction-based scheduling strategy calculates the traffic distribution ratio by predicting the congestion degree or arrival time of each path. Forward prediction scheduling (FPS) [30] can reduce the number of out-of-order packets by predicting the arrival time of each sub-path. In [31], the authors proposed Fine-Grained Forward-Prediction-Based Dynamic Packet Scheduling (F<sup>2</sup>PDPS), which allocates more packets to under-scheduled sub-flows by estimating the number of packets being transmitted by the sub-flows. Additionally, many researchers have proposed using a central controller to collect network topology and path congestion information and allocate data flow to each transmission path. Shafiee et al. [32] proposed a flow-level congestion-aware load balancing method which dynamically adjusts path weights based on bandwidth utilization and prioritizes flow allocation to the path with the minimum weight. In [33], the authors proposed to use a central controller to determine the transmission path for each packet to improve network utilization and reduce queuing delay.

#### 3. System Design

In this section, we introduce a multipath transmission mechanism applied to ICN to improve transmission efficiency and stability by exploiting the characteristics and functions of ICN.

#### Multipath Transmission Mechanism

As the IP network is the main form of the current Internet, it is obviously unrealistic to overthrow the existing network infrastructure to establish a new network architecture. In order to reduce deployment cost, the ICN architecture must be compatible with the existing IP network to fully utilize the IP infrastructure. In our previous research [34], we designed an ICN protocol stack that can achieve progressive development by adding an ID layer above the IP layer in the network layer, thus enabling the protocol stack to have ICN functions and features. In ICN, data, network devices, functions, network services and other network entities are all assigned a globally unique entity ID (EID). The EID

has flat and non-semantic characteristics and can be numerical, a character or any other form of data. To implement the cooperation between an ID and IP, an important network component called a Name Resolution System (NRS) must be introduced in ICN to maintain the mapping relationship between the ID and IP. The NRS is deployed in a distributed manner. At the network layer, ID-based addressing and IP-based routing are decoupled: that is, network devices with the ICN protocol stack can initiate resolution requests to the NRS according to a specific ID to obtain the IP addresses bound to that ID, and the connection is then initiated to a specific ICN server or router based on IP routing rules. By using an EID as an identifier and an IP as a locator, the problems caused by the semantic overload of IP addresses can be completely solved.

Despite the decoupling of addressing and routing at the network layer, the transmission efficiency and robustness of ICN are still limited by the shortest path transmission mode adopted by the underlying IP routing rules. As the number of flows increases, the congestion probability of the default shortest path increases. Once congestion or failure occurs in the default IP path, it is necessary to recalculate and update the routing rules. Typically, the convergence of routing algorithms usually takes tens of seconds or even several minutes. During this period, packet loss, throughput degradation and even session interruption may occur, which seriously affects the quality of experience. Therefore, the rigid shortest path transmission method has difficulty in meeting users' demands for high reliability and high bandwidth, especially in the scenario of massive content distribution in ICN. To improve the reliability and transmission efficiency, we have proposed a prototype system for multipath transmission based on the functions and characteristics of ICN [34]. Since ICN routers have a complete ICN protocol stack as well as caching and computing capabilities, it is a feasible solution to regard them as relay nodes to form relay paths between the two transmission endpoints. RON has experimentally proved that most of the transmission benefits can be obtained by constructing multiple single-hop paths between the sender and receiver [16]. Similarly, we utilize multiple ICN routers to construct multiple single-hop paths at both endpoints to enhance end-to-end transmission efficiency and robustness. Different from overlay routing, the proposed method implements multipath transmission by operating an IP address with an ID at the network layer instead of application layer. Specifically, the process of multipath transmission can be viewed as a service, and the corresponding multipath transmission service ID (MPSID) is generated.

At network initialization, the ICN routers that can provide relay services register their IP addresses and MPSID mapping relationship in an NRS to form a set of candidate relay nodes. For a multipath connection, the data source can initiate the query to the NRS based on the MPSID to obtain the IP addresses of these relay nodes within the set. Then, several appropriate relay nodes can be selected from the set to build multiple relay paths, according to relay node selection strategy. After receiving the data packet sent by the data source, the relay nodes can initiate the query operation to the NRS, according to the receiver device ID, to obtain the destination IP address. Finally, the relay nodes perform network address translation and forward the data packets to the real destination. For security reasons, the IP addresses of the ICN routers inside the network should be hidden from users outside the network. Therefore, the ICN edge router which the data source accesses performs the name resolution operation. Figure 1 shows the principle of the multipath transmission method. In the IP plane, the source node and destination node need to follow the shortest path rules to transmit data through the default IP path (marked by the blue arrow). From the perspective of the ID plane, the ID-based addressing can integrate multiple paths of underlay networks, creating multiple parallel ID paths (marked by the green arrows) between the data source and the receiver to overcome the rigid and cumbersome drawbacks of IP routing protocol.



Figure 1. The principle of the multipath transmission method in ICN.

#### 4. Multipath Data Scheduling

In Section 3, we introduced a prototype system for multipath transmission in ICN, but we did not involve issues such as path correlation analysis and data scheduling. In this section, we focus on how to reduce the correlation of sub-paths and how to distribute data to each sub-path to enhance transmission efficiency and robustness. First, we propose a relay path selection algorithm based on path correlation to minimize the impact of overlapping links. Secondly, we comprehensively calculate the path state value by combining RTT and packet loss rate, and then propose a data-scheduling algorithm based on path state value.

#### 4.1. Path Correlation Analysis

The necessary condition for multipath transmission is to select several suitable relay ICN routers to establish multiple relay paths at both endpoints. During multipath transmission, the paths of the source node and the relay node and the relay node and the destination node comply with the shortest path rule. Some studies [35] have shown that a small number of nodes with a higher betweenness centrality (BC) frequently appear on the end-to-end shortest path in the actual network environment: that is, a small number of relay nodes with high BC can provide optimal routing for most end-to-end communications. The BC of node *i* can reflect the importance of the node in the network topology, which can be calculated as [36]:

$$BC(i) = \sum \frac{\sigma_{st}(i)}{\sigma_{st}} \qquad s, t, i \in V,$$
(1)

where *V* is the set of all nodes in the entire network,  $\sigma_{st}$  represents the shortest path from *s* to *t* and  $\sigma_{st}(i)$  is the shortest path from *s* to *t* passing through *i*. Therefore, in this paper, we select *M* relay ICN routers with a higher BC to construct candidate node set, and their IP addresses and MPSID relations are registered on the NRS. The size of the candidate relay nodes set depends on the scale of the network. Studies have proved that a small set of candidate relay nodes can provide high-quality transmission services for most connections [37]. In the multipath transmission process, the source node can initiate a query operation to the NRS according to the MPSID to obtain the set of candidate relay nodes and select *k* suitable relay nodes to construct *k* end-to-end relay paths. The ID path formed by the relay node is covered on the existing underlay networks, and the data transmission between ICN nodes is still along the IP shortest path. When the ID paths share the same bottleneck link, the transmission efficiency and reliability of the ID paths are usually limited. Once a shared bottleneck link is congested or fails, the overlapping ID paths will be affected. In order to enhance the reliability and robustness of multipath transmission, it is necessary to analyze the path correlation. When selecting relay nodes, we need to consider not only the path correlation between ID paths and the default IP path but also the link overlap among k ID paths. The setting of k affects the gain effect of using relay paths for transmission. Theoretically, the larger k is, the higher the transmission benefit will be. However, due to the existence of overlapping links in an actual network environment, it is unrealistic and inefficient to use too many relay paths. The authors of [38] have shown that most of the transmission benefits can be obtained by

We illustrate the effect of the degree of sub-path overlap on multipath transmission reliability through a simple mathematical model. Assuming a sub-path *path*<sub>i</sub> is composed of *m* link segments, that is,  $path_i = (s_1, s_2, ..., s_m)$ , where the failure probability of the link segment  $s_i$  is  $p_i$ , then the reliability of the  $path_i$  can be expressed as:

selecting three to four relay paths. Therefore, we set k at three in this paper.

$$R(i) = \prod_{j=1}^{m} (1 - p_j)$$
(2)

For the convenience of calculation, we assume that the failure probability of each link segment is equal and unified for p. Then, the reliability of the *path*<sub>i</sub> can be expressed as:

$$R(i) = (1-p)^m.$$
 (3)

The influence of the overlap degree between sub-paths on the joint reliability is considered. Suppose a sub-path *path<sub>j</sub>* consists of *n* link segments, that is, *path<sub>j</sub>* = ( $s_1, s_2, ..., s_n$ ). The number of overlapping link segments of *path<sub>i</sub>* and *path<sub>j</sub>* is *l*, and *l*  $\leq$  *m*, *n*. Then the overlap degree of *path<sub>i</sub>* and *path<sub>i</sub>* can be expressed as:

$$\rho_{ij} = \frac{2l}{m+n} \tag{4}$$

Based on the above assumptions, we can calculate the joint reliability of  $path_i$  and  $path_j$ , which can be expressed as:

$$R(i,j) = (1-p)^{l} (1 - (1 - (1-p)^{m-l})(1 - (1-p)^{n-l})) = (1-p)^{m} + (1-p)^{n} - (1-p)^{(1-\frac{\rho_{ij}}{2})(m+n)}.$$
(5)

According to Formula (5), it can be known that as  $\rho_{ij}$  increases, R(i, j) will decrease, proving that the joint reliability will decrease as the overlap degree of the sub-paths increases. As shown in Figure 2, there are three paths between the data source and the destination (marked by green, red, and blue arrows). Assuming that the failure probability of each link segment is 0.1. according to Formula (3), the reliability of paths 1, 2 and 3 can be expressed as  $R(1) = (1 - 0.1)^4 \approx 0.6561$ ,  $R(2) = (1 - 0.1)^5 \approx 0.5905$  and  $R(3) = (1 - 0.1)^5 \approx 0.5905$ , respectively. Path 1 and path 2 are overlapping, and the number of overlapping links is 1. According to Formula (5), the joint reliability of path 1 and path 2 can be expressed as  $R(1,2) \approx 0.8162$ . Path 2 and path 3 do not overlap, so their joint reliability can be expressed as  $R(1,3) \approx 0.8592$ . It follows that the mutual overlap between sub-paths reduces the joint reliability.



Figure 2. Example of paths overlapping of multipath transmission.

If two sub-paths share a bottleneck link in the underlying physical network, the bottleneck link is usually the "thin waist" of the two sub-paths and is prone to congestion. Therefore, the state changes (e.g., an increase or decrease in the RTT or packet loss rate) of the bottleneck link will directly reflect on the overlapping sub-paths. For instance, an increase in RTT for one sub-path may be associated with an increase in RTT for another sub-path that overlaps it, and vice versa. Based on this, we believe that the behavior between overlapping sub-paths is highly similar. In this paper, we use the statistical correlation of RTT to estimate the correlation between sub-paths. Assuming that  $RTT_i(n)$ , n = 1, 2, ..., N is the RTT sampling value of  $path_i$ , and  $RTT'_i(n)$ , n = 1, 2, ..., N - 1 is the derivative of  $RTT_i(n)$ , which can be expressed as:

$$RTT'_{i}(n) = RTT_{i}(n+1) - RTT_{i}(n), n = 1, 2, \dots, N-1$$
(6)

We introduce a correlation coefficient to describe the overlapping degree between different paths, which is calculated as:

$$Corr(i,j) = \frac{cov(RTT'_{i}(n), RTT'_{j}(n))}{\sqrt{var(RTT'_{i}(n))}\sqrt{var(RTT'_{j}(n))}}$$
(7)

where, Corr(i, j) denotes the correlation coefficient of  $path_i$  and  $path_j$ ,  $cov(RTT'_i(n), RTT'_j(n))$ denotes the covariance of RTT variation of  $path_i$  and  $path_j$  and  $var(RTT'_i(n))$  and  $var(RTT'_j(n))$  denote the variance of RTT variation of  $path_i$  and  $path_j$ , respectively. We set a skewness threshold  $\delta$ . If  $Corr(i, j) > \delta$ , then  $path_i$  and  $path_j$  are correlated; otherwise,  $path_i$  and  $path_j$  are uncorrelated. For a multipath connection, we design a relay node-selection-algorithm based on path correlation (see Algorithm 1), which is described as follows:

Step 1: First, initialize the set of candidate relay nodes *R* based on the result returned by the NRS, initialize the set of selected relay nodes *S* to be empty, and add the default IP path to the set of selected paths *P*.

Step 2: If the source node caches the path correlation information about the destination node, the relay node is selected based on this information, and the relay node selection ends; if the source node does not record any path correlation information about the destination node locally, go to Step 3.

Step 3: The source node randomly selects *k* relay nodes from *R* to construct the initial ID paths and calculates the path correlation coefficient during transmission.

Step 4: Confirm whether the selected ID path has a statistical correlation with the paths in *P* in turn. If the current ID path is not correlated with the paths in *P*, add the relay node to *S* and add the ID path to *P*; otherwise, the current ID path is deleted and an equal number of relay nodes are selected from *R* to rebuild the new ID paths.

Step 5: Repeat Step 4 until *k* relay nodes are found or already traverse all the relay nodes in *R* and record the path correlation information locally to provide a basis for node selection for future connections.

Algorithm 1 Relay-Node-Selection Algorithm
1: <b>Input</b> : <i>G</i> ( <i>V</i> , <i>E</i> ), <i>src</i> , <i>dst</i> , <i>k</i> , MPSID
2: Output: S, P
3: $R = \text{QueryFromNRS}(\text{MPSID})$
4: defaultPath = dijkstra_path( <i>src,ds</i> t)
5: P.add(defaultPath)
6: if hit (src, dst) cache then
7: determine <i>S</i> and <i>P</i> based on the path correlation information
8: else if
9: <b>for</b> $i = 1$ to $M$ <b>do</b>
10: <b>if</b> relayPath( <i>i</i> ) is independent of the path in <i>P</i> <b>then</b>
11: $P.add(relayPath(i))$
12: $S.add(Node(i))$
13: end if
14: <b>if</b> $ S  = k$ then
15: break;
16: end if
17: end for
18: end if
19: <b>return</b> <i>S</i> , <i>P</i>

# 4.2. Data Scheduling

After the relay node selection is completed, multiple uncorrelated sub-paths can be established between the source node and destination node. During multipath transmission, even sequentially sent packets may arrive out of order because the sub-paths may differ in terms of bandwidth, latency, packet loss rate, and so on. In this case, the packets sent later must wait for the packets sent earlier to arrive at the receiver, resulting in packets accumulating and blocking in the receiving buffer. Once the receiving buffer overflows, subsequent packets are forcibly dropped, which, in turn, affects the performance of multipath transmission. Therefore, it is important to design an efficient data-scheduling strategy to maintain traffic load balancing among sub-paths. In the ICN multipath transmission scenario, data scheduling should take into account the following two problems: (1) If there are multiple optional sub-paths, how can the priority of each sub-path be determined? (2) How can the size of the scheduling granularity be determined, and how can the data be distributed to be transferred to each sub-path?

To solve the first problem, the scheduler is required to measure the path state accurately and timely and preferentially assign the traffic to the sub-paths with a better state. In ICN, in addition to the basic routing and forwarding functions, ICN routers also support some specific services, such as a multipath transmission service, caching service and multicast service. Therefore, the traffic characteristics in ICN are more complex and variable, which also increases the difficulty of path state measurement. In this paper, the path state value is calculated by combining the RTT and packet loss rate, which can be expressed as:

$$path\_state_i = \omega_1 \times RTT_i^* + \omega_2 \times P_i^*$$
(8)

where  $RTT_i^*$  and  $P_i^*$  respectively represent the normalized RTT and packet loss rate of the *path<sub>i</sub>*, and the corresponding weights of  $\omega_1$  and  $\omega_2$  satisfy  $\omega_1 + \omega_2 = 1$ . To calculate the path state value, the maximum and minimum standardization of the above indicators is

required. Since both the RTT and packet loss rate belong to negative dimensions, their normalization can be expressed as:

$$RTT_i^* = \frac{max\_RTT - RTT_i}{max\_RTT - \min\_RTT}$$
(9)

$$P_i^* = \frac{max\_P - P_i}{max\_P - \min\_P} \tag{10}$$

where *max\_RTT* and min\_*RTT* denote the maximum and minimum RTT of all sub-paths, respectively, and *max\_P* and min\_*P* represent the maximum and minimum packet loss rate of all sub-paths, respectively.

To solve the second problem, the scheduler is required to determine the appropriate size of scheduling granularity. The scheduling granularity refers to the minimum unit of data scheduling during multipath transmission. According to the size of the scheduling granularity, the existing scheduling methods can be divided into packet-based scheduling methods and cluster-based scheduling methods. The granularity of the packet-based scheduling method is finer and facilitates multipath load balancing, but the scheduler needs to make scheduling decisions for each packet, which increases the burden on system processing. The granularity of the cluster-based scheduling method is coarser, which ensures that the packets within the cluster can be transmitted along the same path and reduces the system processing burden. In ICN, the chunk is the basic data unit for transmission. To ensure that the packets within a chunk can arrive in order, we set the scheduling granularity to the chunk level.

During multipath transmission, a principle needs to be followed: that is, if there are idle paths, the idle paths are preferred; if no idle path exists, the chunks to be transmitted are distributed to each sub-path according to the corresponding scheduling strategy. As described in Algorithm 2, we propose a data-scheduling algorithm that distributes the chunk to be transmitted to the each sub-path based on the path state value. For a multipath connection, the state value of each sub-path is initially set to the same value, and the RTT and packet loss rate of each sub-path are statistically counted to calculate the path state value, which is then used to distinguish the priority of each path. As shown in Figure 3, the first data packet of each chunk is assigned to the sub-path with the highest state value, and subsequent data packets of the same chunk are assigned to the same path. The proposed algorithm integrates the impact of RTT and packet loss rate, thereby improving the sensitivity to path congestion.



Figure 3. The principle of the multipath data scheduling.

Algorithm 2 Data-Scheduling Algorithm	
1: Initialization: $max_RTT = 0$ , $max_P = 0$ , $min_RTT = 0$ , $min_P$	= 0
2: <b>for</b> every scheduling cycle during transmission <b>do</b>	
3: <b>for</b> each sub-path <i>i</i> <b>do</b>	
4: count the $P_i$ and $RTT_i$ of selected path	
5: <b>if</b> $P_i > max_P$ <b>then</b>	
6: $max_P = P_i$	
7: end if	
8: <b>if</b> $P_i < \min_P$ <b>then</b>	
9: $\min_P = P_i$	
10: <b>end if</b>	
11: <b>if</b> $RTT_i > max\_RTT$ <b>then</b>	
12: $max_RTT = RTT_i$	
13: <b>end if</b>	
14: <b>if</b> $RTT_i < \min_R TT$ <b>then</b>	
15: $\min\_RTT = RTT_i$	
16: <b>end if</b>	
17: calculate $RTT_i^*$ and $P_i^*$ according to Formula (9) and (	10)
18: calculate $path_state_i$ according to Formula (8)	
19: end for	
20: distribute the chunk to the sub-path with the best state va	lue
21: end for	

# 5. Evaluation

In this section, we implement our proposal based on NS3, which is a discrete, eventdriven, open-source network simulator [39], and conduct a series of simulation experiments to evaluate the performance of our proposal. First, we evaluate the performance of the data-scheduling algorithm in a simple multipath topology. Then, we comprehensively evaluate the performance of the proposed transmission method in real-world topologies.

#### 5.1. Basic Performance Evaluation

As shown in Figure 4, we first apply our proposal to a simple multipath topology. Source node A and destination node B are connected to edge routers E1 and E2, respectively. There are two disjoint multipaths between E1 and E2, namely, E1-R1-R2-E2 and E1-R3-R4-E2. The default bandwidth, delay and loss rate of path E1-R1-R2-E2 have been marked in the figure, and the default bandwidth, delay, and packet loss rate of path E1-R3-R4-E2 will be set as required in subsequent experiments. The default chunk size is set to 2 MB, the packet size is set to 1250 bytes and each router can accommodate 200 packets. In order to fully evaluate the performance of each method, we compare four data-scheduling strategies, namely, Round-Robin Data Scheduling (denoted as RRDS), Lowest RTT First Data Scheduling (denoted as LRFDS) [26], Forward Delay-based Packet Scheduling (denoted as FDPS) [27] and our proposed Path-State-Based Data Scheduling (denoted as PSDS). RRDS indicates that the chunk to be transferred is distributed to each sub-path in polling mode; LRFDS indicates that the chunk to be transferred is preferentially distributed to the sub-path with the lowest RTT; FDPS indicates that the chunk to be transferred is preferentially distributed to the sub-path with the lowest forward delay.

# 5.1.1. The Effect of Delay on Performance

In this subsection, we set the default bandwidth and packet loss rate of path E1-R3-R4-E2 to 100 Mbps and 0%, respectively. The default packet loss rate is set to 0%, which means that no noise packet loss occurs simply because the router has a buffer overflow. In addition, we set the default delay of path E1-R3-R4-E2 to 5 ms, 10 ms, 15 ms, 20 ms, 25 ms and 30 ms to analyze the transmission performance of the proposed method in scenarios with different link delay differences. We used R1 and R3 to construct two disjoint relay paths between the source and destination nodes and transfer 100 chunks.



Figure 4. Simulation topology 1.

Figure 5 shows the average throughput of the four data-scheduling strategies under different link delay differences. When the delay of path E1-R3-R4-E2 is 5 ms, that is, when the delay of two paths is the same, the average throughput of the four scheduling strategies is similar. When the delay of path E1-R3-R4-E2 is 5 ms, that is, when the delay of two paths is the same, the average throughput of the three scheduling methods is similar. As the delay of path E1-R3-R4-E2 increases, the throughput of RRDS decreases gradually and is always lower than that of the other methods. This is because RRDS mechanically distributes data across sub-paths, ignoring path state differences and resulting in a lower throughput. In addition, compared with LRFDS and FDPS, the proposed PSDS can maintain the highest average throughput. This is because the proposed PSDS can estimate the real state of the path more accurately by integrating the two factors of RTT and packet loss rate.



Figure 5. Average throughput under different link delays.

Figure 6 shows the proportion of out-of-order chunks of four data-scheduling strategies under different link delay differences. As the delay of path E1-R3-R4-E2 increases, the out-of-order chunks of the methods increase gradually. When the delay of two paths is equal, the out-of-order chunks of the four scheduling methods are close. However, with the increase in the delay difference between two paths, the out-of-order chunks of RRDS remain the highest, indicating that the method has difficulty solving the out-of-order problem. In addition, compared with the LRFDS and FDPS, the proposed PSDS can still maintain the proportion at a lower level when the path delay difference is large.



Figure 6. Proportion of out-of-order chunks under different link delays.

# 5.1.2. The Effect of Loss Rate on Performance

In this subsection, we set the default bandwidth and delay of the path E1-R3-R4-E2 to 100 Mbps and 5 ms, respectively. In addition, we set the default packet loss rate of E1-R3-R4-E2 as 0.01%, 0.02%, 0.03%, 0.04%, and 0.05%, respectively, to analyze the impact of packet loss rate differences on transmission performance. Similarly, we have source node A establish a multipath connection with destination node B and transfer 100 chunks.

Figure 7 shows the relationship between the average throughput and packet loss rate differences of the four scheduling strategies. As the packet loss rate of the path E1-R3-R4-E2 increases, the average throughput of the methods decreases gradually, and the throughput of RRDS is always lower than that of the other three methods. FDPS estimates the path state by using the forward delay, thus increasing the throughput compared to LRFDS; however, the improvement is not significant. In addition, compared with the other methods, the proposed PSDS can maintain the highest average throughput. This is because PSDS considers both RTT and packet loss rate so that it can accurately perceive the real path state.



Figure 7. Average throughput under different packet loss rates.

Figure 8 shows the proportion of out-of-order chunks under the different packet loss rates of the four scheduling strategies. As the packet loss rate of path E1-R3-R4-E2 increases, the proportion of the methods increases gradually. This is because a high packet loss rate will result in a large number of retransmissions and a reduction in throughput. Compared with other methods, the proposed PSDS can maintain the proportion of out-of-order chunks at a low level when the path packet loss rate differs greatly.





#### 5.1.3. The Effect of Bandwidth on Performance

In this subsection, we set the default packet loss rate and delay for the path E1-R3-R4-E2 to 0% and 5 ms, respectively. In addition, we set the default bandwidth of E1-R3-R4-E2 to 90 Mbps, 80 Mbps, 60 Mbps, 40 Mbps and 20 Mbps, respectively, to analyze the impact of bandwidth differences on transmission performance. Similarly, we have source node A establish a multipath connection with destination node B and transfer 100 chunks.

Figure 9 shows the relationship between the average throughput and bandwidth differences of the four scheduling strategies. As the available bandwidth of the path E1-R3-R4-E2 decreases, the average throughput of the four methods gradually decreases, and the throughput of RRDS is always lower than that of the other three methods. In addition, compared with the LRFDS and FDPS, the proposed PSDS can maintain the highest average throughput, which indicates that our method performs well in the scenario in which the path bandwidths differ greatly.



Figure 9. Average throughput under different bandwidths.

Figure 10 shows the proportion of the out-of-order chunks of the these four datascheduling strategies in different link bandwidths. As the available bandwidth of path E1-R3-R4-E2 decreases, the proportion of out-of-order chunks of the methods increases gradually. The proportion of RRDS is always kept at a high level. We can also ascertain that the proposed PSDS can maintain a relatively minimum out-of-order proportion when the path bandwidths differ greatly.



Figure 10. Proportion of out-of-order chunks under different bandwidths.

#### 5.2. Comprehensive Evaluation

As shown in Figure 11, to bring the simulation closer to the real network environment, we use two real-world backbone topologies from the dataset provided by Topology Zoo [40], namely, BT North America and GEANT. The default bandwidth between nodes in the topology is uniformly distributed between 25 Mbps and 125 Mbps, the default delay between nodes is set to 1 ms and default loss rate is set to 0%. The default packet size is set to 1250 bytes, and each router can accommodate 200 packets. In addition, we select 10 nodes with the highest BC in these two topologies as relay nodes and randomly select 20 groups of sending and receiving node pairs. Each group of node pairs uses a multipath method to transmit 400 M data, among which the first 200 M of data is used to collect the path correlation information, and the last 200 M of data is used to collect the experimental results. In addition, we set the chunk size to 128 KB, 256 KB, 512 KB, 1 MB and 2 MB to evaluate the performance of each method under different scheduling granularities.



Figure 11. Simulation topology 2: (a) BT North America; (b) GEANT.

Figure 12 shows the average throughput of each group of node pairs under different scheduling granularities in these two topologies. The results in Figure 12a,b are consistent. We can see that the average throughput of the RRDS is always lower than the other methods under different scheduling granularities. This is because RRDS mechanically distributes data to each sub-path in a polling manner, which leads to the problem of an unbalanced load. In addition, compared with the LRFDS and FDPS, our proposed PSDS can maintain the highest average throughput, which also reflects the effectiveness of our proposal.



**Figure 12.** Average throughput under different scheduling granularities: (**a**) BT North America; (**b**) GEANT.

Figure 13 shows the proportion of the out-of-order chunks of the four scheduling strategies under different scheduling granularities in these two topologies. Similarly, the results of Figure 13a,b are consistent. It can be observed that under different scheduling granularities, the proportion of out-of-order chunks of the RRDS is always higher than that of the other three methods. Compared with RRDS, the LRFDS and FDPS alleviate the out-of-order problem to a certain extent by perceiving the path state through the RTT and forward delay. Additionally, our proposed PSDS is able to maintain a relatively low level of out-of-order proportion compared to other methods, indicating that our proposal can achieve significant performance in real network topologies.



**Figure 13.** Proportion of out-of-order chunks under different scheduling granularities: (**a**) BT North America; (**b**) GEANT.

# 6. Conclusions

In this paper, we design and evaluate a multipath data-transmission-scheduling strategy based on path correlation applied in ICN, aiming to fully utilize network bandwidth resources and realize load balancing. We first introduce the principle of the multipath transmission mechanism, which combines the functions and characteristics of ICN and is able to utilize the infrastructure of IP network. Then, we estimate the path correlation according to the statistical correlation of RTT and propose a relay node selection strategy; In addition, we design a multipath data-scheduling strategy by combining the RTT and packet loss rate to ensure the load balance. Finally, we evaluate the proposed method through simulation experiments. The results show that the proposed data-scheduling strategy can maintain high bandwidth utilization while alleviating the packet out-of-order problem.

Future work should focus on two aspects. Firstly, due to the differences between simulation and real implementation, future work will consider the deployment of our proposal in real network environments. Secondly, given that existing research still lacks a reasonable multipath congestion control mechanism, future work will mainly focus on this aspect to avoid congestion or failure caused by link overload.

**Author Contributions:** Conceptualization, Y.X., H.N. and X.Z.; methodology, Y.X., H.N. and X.Z.; software, Y.X.; writing—original draft preparation, Y.X.; writing—review and editing, H.N. and X.Z.; supervision, X.Z.; project administration, X.Z.; funding acquisition, H.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Strategic Leadership Project of the Chinese Academy of Sciences: SEANET Technology Standardization Research System Development (Project No. XDC02070100).

Data Availability Statement: Not applicable.

Acknowledgments: We would like to express our gratitude to Jinlin Wang and Zhiyuan Wang for their meaningful support of this work.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Passarella, A.J.C.C. A survey on content-centric technologies for the current Internet: CDN and P2P solutions. *Comput. Commun.* 2012, 35, 1–32. [CrossRef]
- Ahlgren, B.; Dannewitz, C.; Imbrenda, C.; Kutscher, D.; Ohlman, B.J.I.C.M. A survey of information-centric networking. *IEEE Commun. Mag.* 2012, 50, 26–36. [CrossRef]
- Jacobson, V.; Smetters, D.K.; Thornton, J.D.; Plass, M.F.; Briggs, N.H.; Braynard, R.L. Networking named content. In Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies, Rome, Italy, 1–4 December 2009; pp. 1–12.
- Zhang, L.; Afanasyev, A.; Burke, J.; Jacobson, V.; Claffy, K.; Crowley, P.; Papadopoulos, C.; Wang, L.; Zhang, B.J.A.S.C.C.R. Named data networking. ACM SIGCOMM Comput. Commun. Rev. 2014, 44, 66–73. [CrossRef]
- Saxena, D.; Raychoudhury, V.; Suri, N.; Becker, C.; Cao, J.J.C.S.R. Named data networking: A survey. *Comput. Sci. Rev.* 2016, 19, 15–55. [CrossRef]
- Trossen, D.; Parisis, G.J.I.C.M. Designing and Realizing an Information-Centric Internet. *IEEE Commun. Mag.* 2012, 50, 60–67. [CrossRef]
- Raychaudhuri, D.; Nagaraja, K.; Venkataramani, A.J.A.S.M.C.; Review, C. Mobilityfirst: A robust and trustworthy mobility-centric architecture for the future internet. ACM SIGMOBILE Mob. Comput. Commun. Rev. 2012, 16, 2–13. [CrossRef]
- Dannewitz, C.; Kutscher, D.; Ohlman, B.; Farrell, S.; Ahlgren, B.; Karl, H.J.C.C. Network of Information (NetInf)—An informationcentric networking architecture. *Comput. Commun.* 2013, *36*, 721–735. [CrossRef]
- 9. Wang, J.; Chen, G.; You, J.; Sun, P. SEANet: Architecture and Technologies of an On-Site, Elastic, Autonomous Network. J. Netw. New Media 2020, 9, 1–8.
- 10. Hopps, C. Analysis of an Equal-Cost Multi-Path Algorithm. Available online: https://datatracker.ietf.org/doc/html/rfc2992 (accessed on 9 February 2023).
- Filsfils, C.; Previdi, S.; Ginsberg, L.; Decraene, B.; Litkowski, S.; Shakir, R. Segment Routing Architecture. Available online: https://www.rfc-editor.org/rfc/rfc8402 (accessed on 9 February 2023).
- 12. Ventre, P.L.; Tajiki, M.M.; Salsano, S.; Filsfils, C. SDN architecture and southbound APIs for IPv6 segment routing enabled wide area networks. *IEEE Trans. Netw. Serv.* 2018, 15, 1378–1392. [CrossRef]
- Scharf, M.; Ford, A. Multipath TCP (MPTCP) Application Interface Considerations. Available online: https://www.rfc-editor. org/rfc/rfc6897 (accessed on 9 February 2023).
- 14. Stewart, R.; Metz, C. SCTP: New transport protocol for TCP/IP. IEEE Internet Comput. 2001, 5, 64–69. [CrossRef]
- 15. Savage, S.; Anderson, T.; Aggarwal, A.; Becker, D.; Cardwell, N.; Collins, A.; Hoffman, E.; Snell, J.; Vahdat, A.; Voelker, G.J.I.M. Detour: Informed Internet routing and transport. *IEEE Micro* **1999**, *19*, 50–59. [CrossRef]
- Andersen, D.; Balakrishnan, H.; Kaashoek, F.; Morris, R. Resilient overlay networks. In Proceedings of the Eighteenth ACM Symposium on Operating Systems Principles, Banff Alberta, AB, Canada, 21–24 October 2001; pp. 131–145.
- 17. Luo, J.; Hu, J.; Wu, D.; Li, R. Opportunistic routing algorithm for relay node selection in wireless sensor networks. *IEEE Trans. Ind. Inform.* **2014**, *11*, 112–121. [CrossRef]

- Tuah, N.; Ismail, M. Extending lifetime of heterogenous wireless sensor network using relay node selection. In Proceedings of the 2013 International Conference of Information and Communication Technology (ICoICT), Bandung, Indonesia, 20–22 March 2013; pp. 17–21.
- Romanov, A.; Myachin, N.; Sukhov, A. Fault-tolerant routing in networks-on-chip using self-organizing routing algorithms. In Proceedings of the IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 13–16 October 2021; pp. 1–6.
- Zhou, J.; Tewari, M.; Zhu, M.; Kabbani, A.; Poutievski, L.; Singh, A.; Vahdat, A. WCMP: Weighted cost multipathing for improved fairness in data centers. In Proceedings of the Ninth European Conference on Computer Systems, Amsterdam, The Netherlands, 14–16 April 2014; pp. 1–14.
- 21. Moy, J. OSPF Version 2. Available online: https://www.rfc-editor.org/rfc/rfc2178 (accessed on 9 February 2023).
- Han, J.; Watson, D.; Jahanian, F. Topology aware overlay networks. In Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, FL, USA, 13–17 March 2005; pp. 2554–2565.
- Jin, J.; Nahrstedt, K. QoS service routing in one-to-one and one-to-many scenarios in next-generation service-oriented networks. In Proceedings of the IEEE International Conference on Performance, Computing, and Communications, Phoenix, AZ, USA, 15–17 April 2004; pp. 503–510.
- Roy, S.; Pucha, H.; Zhang, Z.; Hu, Y.C.; Qiu, L. On the placement of infrastructure overlay nodes. *IEEE/ACM Trans. Netw.* 2009, 17, 1298–1311. [CrossRef]
- Cha, M.; Moon, S.; Park, C.-D.; Shaikh, A. Placing relay nodes for intra-domain path diversity. In Proceedings of the 25th IEEE International Conference on Computer Communications, Barcelona, Spain, 23–29 April 2006; pp. 1–12.
- Paasch, C.; Ferlin, S.; Alay, O.; Bonaventure, O. Experimental evaluation of multipath TCP schedulers. In Proceedings of the 2014 ACM SIGCOMM Workshop on Capacity Sharing Workshop, Chicago, IL, USA, 18 August 2014; pp. 27–32.
- 27. Le, T.-A.; Bui, L.X. Forward delay-based packet scheduling algorithm for multipath TCP. *Mob. Netw. Appl.* **2018**, 23, 4–12. [CrossRef]
- Vu, V.A.; Walker, B. Redundant multipath-tcp scheduling with desired packet latency. In Proceedings of the 14th Workshop on Challenged Networks, Los Cabos, Mexico, 21–25 October 2019; pp. 7–12.
- Cao, Z.; Wang, Z.; Zegura, E. Performance of hashing-based schemes for internet load balancing. In Proceedings of the IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, Israel, 26–30 March 2000; pp. 332–341.
- Mirani, F.H.; Boukhatem, N.; Tran, M.A. A data-scheduling mechanism for multi-homed mobile terminals with disparate link latencies. In Proceedings of the 2010 IEEE 72nd Vehicular Technology Conference-Fall, Ottawa, ON, Canada, 6–9 September 2010; pp. 1–5.
- Ni, D.; Xue, K.; Hong, P.; Shen, S. Fine-grained forward prediction based dynamic packet scheduling mechanism for multipath TCP in lossy networks. In Proceedings of the 2014 23rd International Conference on Computer Communication and Networks (ICCCN), Shanghai, China, 4–7 August 2014; pp. 1–7.
- 32. Shafiee, M.; Ghaderi, J. A simple congestion-aware algorithm for load balancing in datacenter networks. *IEEE/ACM Trans. Netw.* **2017**, 25, 3670–3682. [CrossRef]
- Perry, J.; Ousterhout, A.; Balakrishnan, H.; Shah, D.; Fugal, H. Fastpass: A centralized" zero-queue" datacenter network. In Proceedings of the 2014 ACM Conference on SIGCOMM, Chicago, IL, USA, 17–22 August 2014; pp. 307–318.
- 34. Xu, Y.; Ni, H.; Zhu, X. An Effective Transmission Scheme Based on Early Congestion Detection for Information-Centric Network. *Electronics* **2021**, *10*, 2205. [CrossRef]
- 35. Liao, J.; Tian, S.; Wang, J.; Li, T.; Qi, Q. Load-balanced one-hop overlay multipath routing with path diversity. *KSII Trans. Internet Inf. Syst.* **2014**, *8*, 443–461.
- Brandes, U. On variants of shortest-path betweenness centrality and their generic computation. Soc. Netw. 2008, 30, 136–145. [CrossRef]
- Cohen, R.; Raz, D. Cost-effective resource allocation of overlay routing relay nodes. *IEEE/ACM Trans. Netw.* 2013, 22, 636–646. [CrossRef]
- Bui, V.; Zhu, W.; Bui, L. Optimal relay placement for maximizing path diversity in multipath overlay networks. In Proceedings of the 2008 IEEE Global Telecommunications Conference, New Orleans, LA, USA, 30 November–4 December 2008; pp. 1–6.
- 39. ns-3 Network Simulator. Available online: https://www.nsnam.org (accessed on 1 April 2023).
- 40. Knight, S.; Nguyen, H.X.; Falkner, N.; Bowden, R.; Roughan, M. The internet topology zoo. *IEEE J. Sel. Areas Commun.* 2011, 29, 1765–1775. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.