

## Research Article

# Micro Operator Design Pattern in 5G SDN/NFV Network

Chia-Wei Tseng <sup>1</sup>, Yu-Kai Huang <sup>1</sup>, Fan-Hsun Tseng <sup>2</sup>, Yao-Tsung Yang <sup>1</sup>,  
Chien-Chang Liu <sup>1</sup> and Li-Der Chou <sup>1</sup>

<sup>1</sup>Department of Computer Science and Information Engineering, National Central University, Taoyuan 32001, Taiwan

<sup>2</sup>Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taipei 10610, Taiwan

Correspondence should be addressed to Li-Der Chou; [cld@csie.ncu.edu.tw](mailto:cld@csie.ncu.edu.tw)

Received 27 February 2018; Accepted 30 May 2018; Published 10 July 2018

Academic Editor: Shao-Yu Lien

Copyright © 2018 Chia-Wei Tseng et al. 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.

The trend of 5G mobile networks is increasing with the number of users and the transmission rate. Many operators are turning to small cell and indoor coverage of telecom network service. With the emerging Software Defined Networking and Network Function Virtualization technologies, Internet Service Provider is able to deploy their networks more flexibly and dynamically. In addition to the change of the wireless mobile network deployment model, it also drives the development trend of the Micro Operator ( $\mu$ O). Telecom operators can provide regional network services through public buildings, shopping malls, or industrial sites. In addition, localized network services are provided and bandwidth consumption is reduced. The distributed architecture of  $\mu$ O tackles computing requirements for applications, data, and services from cloud data center to edge network devices or to the micro data center of  $\mu$ O. The service model of  $\mu$ O is capable of reducing network latency in response to the low-latency applications for future 5G edge computing environment. This paper addresses the design pattern of 5G micro operator and proposes a Decision Tree Based Flow Redirection (DTBFR) mechanism to redirect the traffic flows to neighbor service nodes. The DTBFR mechanism allows different  $\mu$ O's to share network resources and speed up the development of edge computing in the future.

## 1. Introduction

Many new operators have chosen to turn to small cells and micro telecommunication network service that covers indoor areas in the wake of the coming of the 5G era, the changes of habits of users when going online, and the increasing demands for applications. Telecom operators can provide services via a variety of access networks such as 3G, 4G, 5G, and even Wi-Fi [1–3]. With the emerging Software Defined Networking (SDN)/Network Function Virtualization (NFV) technologies, Internet Service Provider (ISP) is able to deploy their networks more flexibly and dynamically [4–6]. ISPs are now capable of providing localized services by means of public buildings, shopping malls, or industrial facilities, and this has not only changed the model of distributing and establishing mobile network but also gave birth to the idea of the micro operator ( $\mu$ O)[7–9]. Factors such as scarcity of spectrum resources, the effective employment of bandwidth, the scale of the mobile market, and consumers' options for the diversity of services conspire to produce the

inception of Mobile Virtual Network Operator (MVNO), which in turn brings more diverse applications and services to the telecommunication market [10]. MVNO can refer to any telecommunication operators who provide wireless communication services to consumers and do not have their own wireless network infrastructures. In order to expand the coverage of their business, Mobile Networker Operators (MNO) might choose to cooperate with other MNOs by renting the bandwidth and time of use in their domains to rapidly develop the business of MVNO. MVNO provides brand new opportunities for development for the matured mobile communication market and spurs the telecommunication market to move toward a service-oriented competition. As for 5G mobile network operators, development of small cells and micro telecommunication network service that covers indoor areas is now being initiated to satisfy users' demand for indoor coverage, maturing the concept of  $\mu$ O [11–13].

Figure 1 shows the  $\mu$ O network architecture.  $\mu$ O is relatively small in scale and holds limited resources to provide particular and necessary services to a certain number of

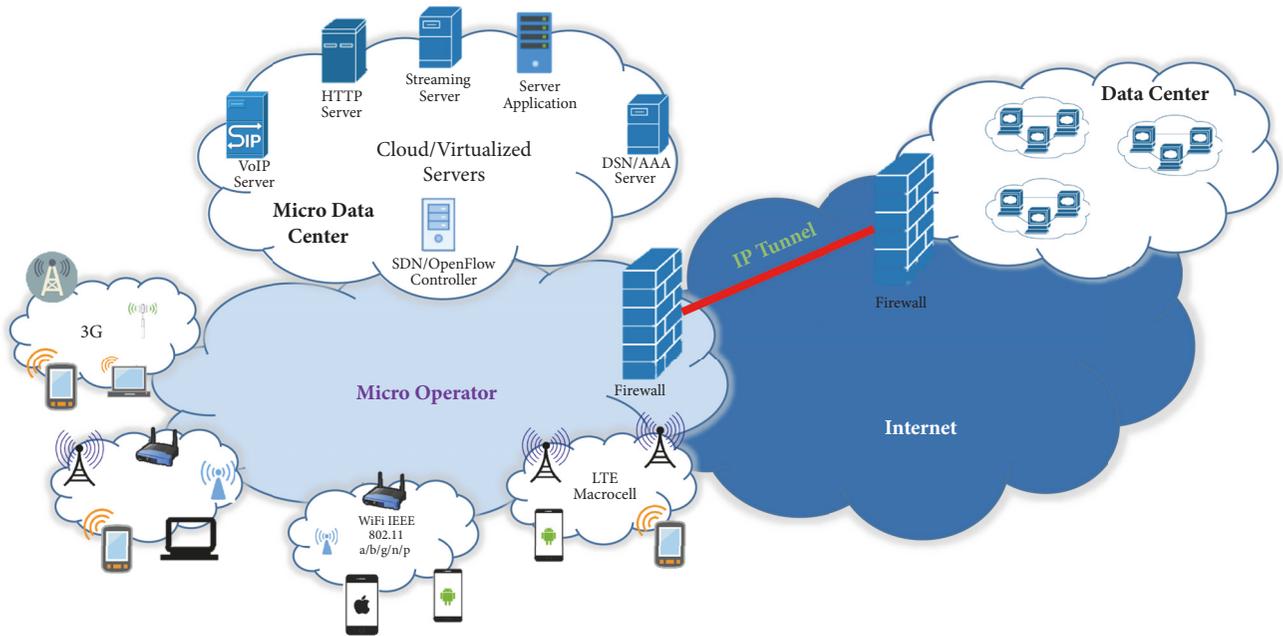


FIGURE 1: Micro operator network architecture.

users. The  $\mu O$  is originated from Oulu University in Finland that consists of three features [14]. (1) Mobile devices access local network specific service (2) spatially confined area with specific service (3) dependent on appropriate available network resources. Facilities like hospitals, schools, large conference, sports venues, shopping malls, hypermarkets, and factories can use local area network service to satisfy users' needs for all sorts of application services.

Restricted by resources and space,  $\mu O$  has the character of regionalization service that allows it to provide, under circumstances where hardware infrastructures are scarce and the resources are limited, different regional services according to different domains, making it possible for users of mobile devices to gain access to services in the nearby regions. Apart from reducing the consumption of bandwidth resources by offering nearby network services, this kind of localization service can also transfer applications, data, and the computing of services from nodes in the data center on cloud to edge nodes in logical LAN to be processed and realize the developing environment for edge computing [15, 16], reduce network latency, and satisfy 5G demands for amelioration of latency.

The purpose of this paper is to discuss the  $\mu O$  design pattern in 5G SDN/NFV network and establishes an experimental environment to evaluate the decision tree based traffic flow redirection mechanism. To realize the customized experiment environment, the network functions virtualization is taken into consideration in our research. Our contributions are listed as follows:

- (1) A  $\mu O$  design pattern is constructed based on SDN and NFV that combines network slicing and tunneling technologies to build a network infrastructure for  $\mu O$ s.

- (2) A DTBFR mechanism for  $\mu O$  is proposed and the decision tree theory is utilized to serve as the reference in the determining the SDN traffic flows direction.
- (3) We established a  $\mu O$  environment for validation of the DTBFR mechanism and  $\mu O$ s communication experiments.

The rest of the paper is organized as follows: in Section 2, the background and related works are addressed. Section 3 describes the  $\mu O$  design pattern and DTBFR mechanism. Section 4 presents the experiment results. The last section concludes this paper.

## 2. Background and Related Works

The development of 5G application services will be largely on the Internet of Things (IoT) and encourages the new communication market to move toward a more vertical-subdivision one [17, 18]. The result will be the formation of different emerging application service scenarios and more diverse demands for networks. However, as far as telecommunication is concerned, despite the fact that there is now a globally agreed requirement on IoT constructed by 5G on characters such as transmitting speed, capacity, coverage, and security, there is still room for the business model of 5G to improve itself. 5G mobile broadband network puts the accent on small cells/base stations, how to amplify indoor coverage, providing faster services for users, and reducing the latency in network transmission, all of which pose a great challenge for telecommunication operators [19–21]. Furthermore, the appearance of MVNOs has brought new opportunities for development for MNOs who have not yet gained licenses to operate mobile communication business.

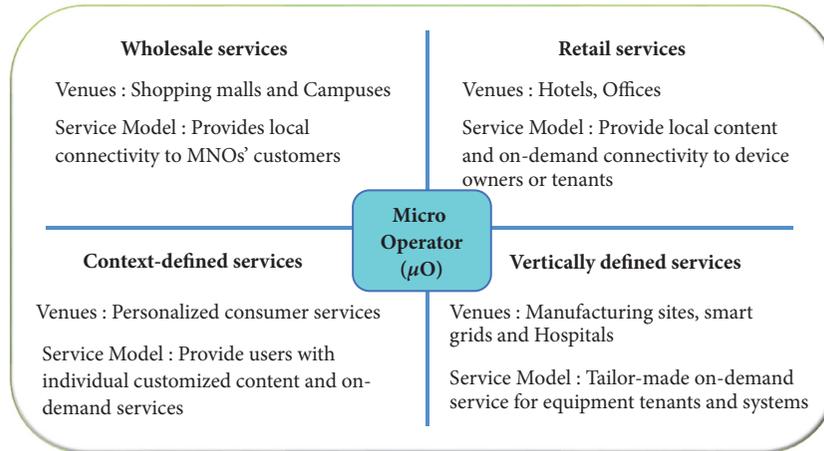


FIGURE 2: Micro operator service model [24].

MVNOs use MNOs' spectrum and network to provide customized mobile communication services, the mobile virtual private network for enterprises, or scores of other micro markets that operators still have not stretched their services to such as less lucrative or regional local emerging markets. A trend of regionalization has been observed among small cell/base stations and the regional services do not possess the characters of public area network, industrial application network, and a hybrid network.  $\mu O$  is an emerging network business model that has started developing under this backdrop [22]. Reference [14] illustrates the relation between  $\mu O$  and MNO and scrutinizes the use cases of  $\mu O$ -MNO to set the foundation for the possible business model for  $\mu O$ . Reference [23] proposes a mechanism in regard to  $\mu O$ 's shared spectrum access communication and shares the infrastructure of mono-network/physical network via virtual technology to fully utilize the valuable bandwidth resource.

As shown in Figure 2, the service model of  $\mu O$  can be divided into four quadrants:

- (1) The first quadrant is closed, controlled, independent, and general service that often takes place in office. The users in this region can gain access to the services provided by the region.
- (2) The second quadrant is shared, transparent, and general service that often takes place in shopping malls and schools. Such regions are often packed with people and a great deal of data is required as well. A micro data center might be set in the region for users to quickly gain access to the services.
- (3) The third quadrant is shared, transparent, and special service that often takes place in scenarios like telematics network and customized services. Such services are likely to be used by end-users. The communication among end-users requires special network protocols.
- (4) The fourth quadrant is closed, controlled, independent, and special service that might take place in hospitals, automated factories, and smart grids. Hospitals provide irreplaceable medical services. Users

will have to connect to such regions before gaining access to the services.

Due to the increasing demand of mobile device users accessing Internet services, operators need to dynamically adjust and combine to meet the requirements of different applications in order to improve network performance. The service architecture of 5G  $\mu O$  must also be developed in conjunction with different technologies such as SDN and NFV [25, 26]. SDN abstracts the network architecture by decoupling the network control and forwarding functions enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services [27]. By separating the control plane from the data forwarding plane and virtualizing all of the connections, administrator can remove the hard-wired barriers of networks and quickly change structures to suit their needs. In addition to the SDN, NFV is a core structural change in the way telecommunication infrastructure gets deployed. The goal of NFV is to enable service providers to reduce costs and faster service delivery. The requirements and open standards that underpin NFV are being developed by the European Telecommunications Standards Institute (ETSI) [28].

Since 5G technology is now undergoing rapid development, the conventional network infrastructure is not able to meet 5G's diverse demands and  $\mu O$ 's development of service model in the future. Therefore, the international standards organization and equipment manufacturers now set out to promote the technology of network slicing which allows ISP to use SDN and NFV technologies to realize network virtual slicing, the division of several different service scenarios, and the provision of customized network service [29–31]. The 5G white paper proposed by Next Generation Mobile Networks (NGMN) that expounds 5G's network infrastructure in the future had included the concept of network slicing technology in order to support particular connection forms and use particular methods to deal with the communication services of control/user-plane [32]. To satisfy the requirements of 5G network's flexible configuration, NOKIA had come up with the programmable 5G network infrastructure where network slices of several independent virtual subnetworks

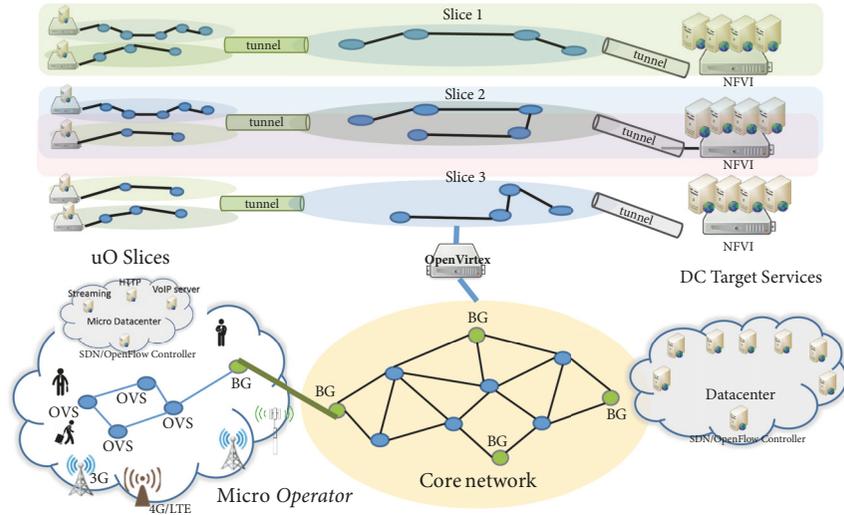


FIGURE 3: A  $\mu$ O design pattern with network slicing.

are constructed in the same basic infrastructure to serve the purpose of the specific application [33]. Network slicing technology can divide network resources into multiple independent virtual networks. Not only can it satisfy the diverse demands of small cells/base stations and  $\mu$ O but it can also work as the crucial technology for the development of 5G network communication. As for virtualization tools, in order to realize the function of network slicing, [34] proposes the concept of Network Virtualization Layer–FlowVisor that divides a physical network into several virtual networks that each possesses their own controllers and enables personal virtual network to be built in the same physical network equipment to achieve the goal of resource allocation. Reference [35] proposes a platform of network virtualization called OpenVirtex (OVX) that is capable of realizing network virtualization for multiple tenants. OVX can map a user's network topology to a physical topology according to the user's needs and accomplish the network-connecting process. The major difference between OVX and FlowVisor is that users of OVX do not have to worry about the basic physical topology and can define their own virtual network topology. OVX will assist its users to accomplish the mapping of virtual network topology to physical network topology.

The conventional offline benchmark planning is no longer suitable to satisfy 5G  $\mu$ Os' efficient utilization of resources and demands of sharing and the full utilization of computing resources to respond to the complicated situations of different demands to determine the direction of the traffic flow. Machine learning is on the list of the most potential candidates [36]. It provides a much stronger and complicated decision-making ability. Decision tree is one of the most popular machine learning algorithms used all along. Decision tree, on the other hand, is highly efficient supervised learning model that is ideal for the prediction of classification and regression data type such as ID3 and C4.5 [37, 38]. C4.5 is the refined version of feature selection in ID3 algorithm that can be used to process continual class label and the problem of missing data in training data. It can also correct

the problem of overfitting of the decision tree by means of pruning. As for the application of route choosing and traffic flow identification in regard of mobile edge network, [39] uses the characters of SDN and combines them with decision tree algorithm to categorize traffic flow, allowing the controller to choose an access point from the local network that is suitable to be connected to a mobile device according to the level of congestion of the backhaul route. Reference[40] proposes a model called AMPS that combines SDN and machine learning. The model is capable of categorizing the differences of each traffic flow by means of learning packet identification and then chooses the best route of transmission for the traffic flow to achieve the automation of the optimal selection of routing path.

### 3. 5G $\mu$ O Design Pattern and DTBFR Mechanism

The interaction between SDN and NFV enables the 5G network to build a system infrastructure in an abstract manner and further increase the flexibility of the network, allows the vertical system to be divided into multiple sliced constructor blocks, and builds a network that is connectable, programmed, and virtual. The operator can advance to use network slicing technology to virtualize networks and, as the aforesaid paragraphs, flexibly divide a physical network into multiple independent and segregated  $\mu$ O networks according to different usage scenarios. As for the design of the infrastructure, this paper uses SDN and NFV technologies as the base and combines network slicing and tunneling technologies to build a network infrastructure for  $\mu$ Os. The infrastructure allows users of different  $\mu$ Os to be concatenated via tunneling technology and then realizes the rapid connection of network to effectively enhance the interconnectivity of networks.

The proposed  $\mu$ O design pattern is shown in Figure 3. Network slicing technology realizes the logical partition of  $\mu$ Os' networks via OpenVirtex. The connection between the

core network and the  $\mu$ O is carried out by the tunnel built by SDN border gateway (BG). The Internet cloud data center and the micro data center are constructed by NFV infrastructure defined by ETSI to save the cost of equipment investment. The mission of SDN controller is to use OpenFlow to construct a passageway between BG at the edge of the core network and BG at the edge of the  $\mu$ O's network. Once the SDN controller begins executing the matching and concatenation between network passageways, the  $\mu$ O's network can build a connection via the tunnel and the main core network. The  $\mu$ O can continue finishing the construction of virtual network via OpenVirtex's virtual slicing technology, allowing users of the  $\mu$ O to gain access to the data on the nearby micro data center. The users can also, via tunnels, connect to the cloud data center on the Internet to access the network service of particular application. The proposed architecture pattern combines network slicing and tunneling to realize the communication model for  $\mu$ O's and further integrates bandwidth controlling technology to be applied to the communication bandwidth application of the  $\mu$ O's network. This will increase the utilization ratio of network resources and the efficiency of traffic flow and result in a better quality of experience (QoE) for the network users.

In response to the demand of  $\mu$ O's network resource distribution that allows users to gain access of nearby network resources, the paper proposes a DTBFR mechanism for  $\mu$ O and uses decision tree theory to serve as the reference in the determining the SDN traffic flows direction. Decision tree algorithms have been used for solving predictive analytics problems in the past few years. Decision tree is a supervised machine learning model with simple process intuition and high execution efficiency. Decision tree algorithm is mainly used to conduct systematic result and integration of dataset and to find the special class and label relation during the decision-making process. Compared with other ML models, execution speed is a major advantage. To quantify uncertain information and dataset, the paper uses entropy as a method to gauge uncertainty and level of chaos in (1). Entropy is a measure of the impurity in a collection of training examples. The entropy increases with the increase in uncertainty or randomness and decreases with a decrease in uncertainty or randomness. The production of information comes with uncertainty and entropy can gauge it according to its probability of occurrence [41]. With higher probability, the chaos is more likely to occur and the uncertainty is low, while on the contrary, the uncertainty is high.

$$Entropy(s) = \sum_{i=1}^c P_i \log_2 P_i \quad (1)$$

Decision trees algorithm uses feature selection to guide the decision of the most useful attributes. Different feature selection criteria result in different types of decision trees. ID3 uses entropy and information gain to construct a decision tree. Information gain is used to decide which feature to split on at each step in building the tree. Information gain measures the expected reduction in entropy by partitioning the examples according to an attribute. Calculations of ID3 information gain are shown in (2). The information gains

criteria to measure the strength of the association between an attribute and class. Symbol S is the target class and Symbol A represents the class label. C4.5 is an extension of ID3 which is a similar tree generation algorithm. C4.5 uses gain ratio by splitting the training sets based on its test attributes. The gain ratio takes number and size of branches into account when choosing an attribute. It corrects the information gain by taking the intrinsic information of a split into account. The split information can be defined as shown in (3). The gain ration can be defined as shown in (4). The flowchart of decision tree construction is shown in Figure 4.

$$Gain(S,A) = Entropy(S) - \sum_{j=1}^v \frac{|S_j|}{|S|} Entropy(S_j) \quad (2)$$

$$SplitInfo(S,A) = - \sum_{i=1}^c \frac{|S_i|}{|S|} \log_2 \left( \frac{|S_i|}{|S|} \right) \quad (3)$$

$$GainRatio(S,A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (4)$$

The dataset used by the decision tree consists of both class labels and decision result. Training data collects user application traffic primarily through SDN controller. The dataset includes user's network area, server's CPU/memory, and traffic oriented result. The data storage format is represented as [*user\_area*, *s1\_cpu*, *s2\_cpu*, *s3\_cpu*, *s1\_mem*, *s2\_mem*, *s3\_mem*, *result*]. For example, the dataset [2, 49.0, 35.6, 6.0, 29.0, 13.0, 36.3, 2] corresponds to the class label [ $\mu$ O, *s0\_cpu*, *s0\_memory*, *s1\_cpu*, *s1\_mem*, *s2\_cpu*, *s2\_mem*, *result*]. Training data is a certain percentage of an overall dataset along with a testing set. The more complete training data makes classification easier. The duplicate values and divorced values must be removed first from the dataset. Duplicate values refer the same values for each class label in the dataset. Divorced values are calculated from the average value of the class label. When the value of the class label is greater than the average of  $\pm 5$ , the dataset will be removed from dataset.

After completing training on data collection, we can start to establish the node of the decision tree. Calculate the *Gain(S, A)*, *SplitInfo(S, A)*, and *Gain\_ratio(S, A)* from the class labels of the dataset. Pick the best *Gain\_ratio(S, A)* from the list and save it to BestFeature as the root of the decision tree. The remaining unselected class labels are added to Unselected\_labels. The execution of the recursion begins when the decision tree node is established. Pruning is a technique in machine learning that reduces the size of decision trees by pruning the tree based on the statistical confidence estimates. When the decision tree is created, load the training data and start traverse tree node from root to leaf. We adopted the pessimistic pruning strategy proposed used in C4.5 to avoid the need of pruning set and continuity correction to error rate at each node.

In this paper, the calculation of continuous class labels and discrete class labels is different. Discrete data is the type of data that has clear spaces between values while continuous

TABLE 1: Example of continuous class labels.

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
CPU	85.4	90.3	78.4	80.5	80.5	75.2	66.0	90.3	75.2	80.5	75.8	90.3	75.2	90.3
S	No	No	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes

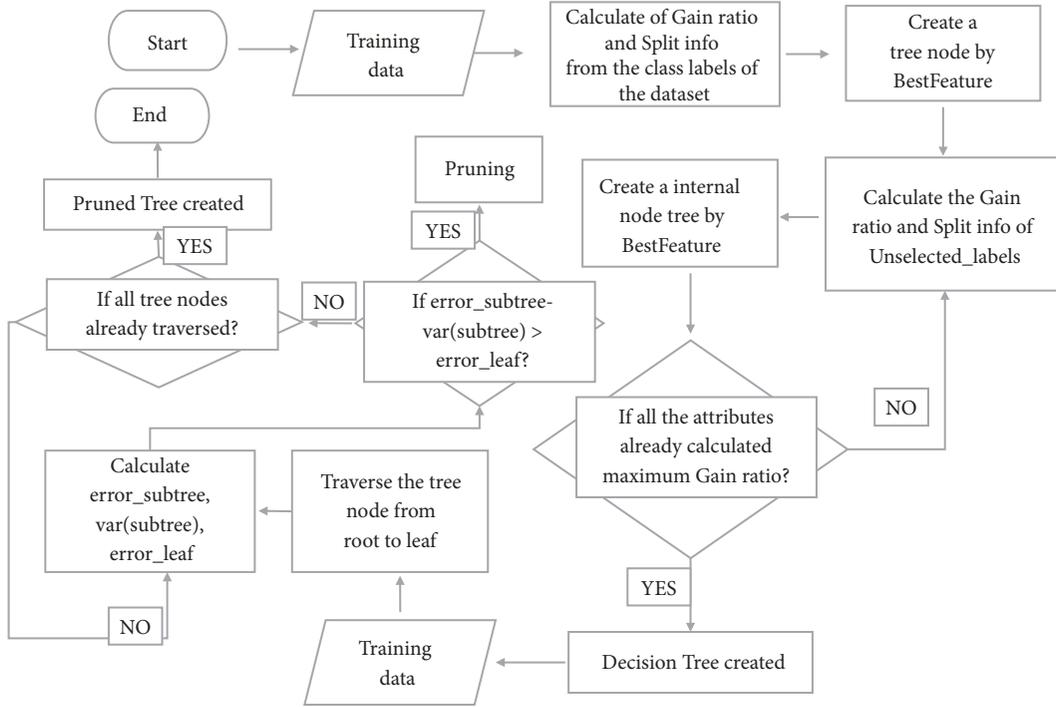


FIGURE 4: The decision tree construction flowchart.

data is data that falls on a continuous sequence. Table 1 shows the example of continuous class label. The second column shows the servers' CPU usage in  $\mu\text{O}$ . The third column S represents the output of the decision. In this example we start with number 66 to calculate the gain ratio. The dataset greater than 66.0 and the dataset less than or equal to 66.0 use (1) to compute the entropy value as 0 and 0.96, respectively. Then, the results are calculated by (2) and (3). Dividing the *Gain* (S, 66.0) by the (3) yields the gain ratio value of 0.129. Table 2 shows the calculation results and the class label CPU 66.0 gets the maximum gain ratio value in this example. Based on the above calculation, the best value is chosen in each round to serve as the node of the tree until the decision tree is established.

Based on the results of Table 2, the decision tree is constructed as shown in Figure 5. In order to correct the overlearning problem caused by the impure information in the decision tree, we adopted the pessimistic error pruning strategy to recursively estimate the error rate of the sample nodes covered by each internal node. Pessimistic error pruning is based on error estimates derived from the training data.

If the internal node's error rate is lower than the estimated value, it will be replaced by the largest number of leaf nodes in the subtree.

Assume that an internal node of the tree covers  $N$  samples with  $E$  errors; then the error rate of the leaf nodes is given by (5). This penalty factor is 0.5. Pessimistic error pruning adds a constant to the training error of a subtree by assuming that each leaf automatically classifies a certain fraction of an instance incorrectly. This fraction is taken to be  $1/2$  divided by the total number of instances covered by the leaf and is called a continuity correction in statistics.

$$E(\text{leaf\_error\_rate}) = \frac{(E + 0.5)}{N} \quad (5)$$

$$E(\text{subtree\_error\_rate}) = \frac{(E + 0.5 * \text{leaf})}{N} \quad (6)$$

After pruning, the internal node becomes a leaf node and the false positive number also need to add a penalty factor. The error rate of subtree is given by (6). Equations (7) and (8) show the false positive number and the standard deviation of

TABLE 2: Calculation result of the continuous class labels.

Value	66.0		75.2		75.8		78.4		80.5		85.4		90.3	
interval	$\leq$	$>$	$\leq$	$>$	$\leq$	$>$	$\leq$	$>$	$\leq$	$>$	$\leq$	$>$	$\leq$	$>$
Yes	1	8	3	6	4	5	5	4	7	2	7	2	9	0
No	0	5	1	4	1	4	1	4	2	3	3	2	5	0
Entropy	0	0.96	0.81	0.97	0.72	0.99	0.65	1	0.76	0.97	0.88	1	0.65	0
Info(S,A)	0.891		0.92		0.89		0.85		0.838		0.91		0.651	
Gain	0.048		0.015		0.045		0.09		0.102		0.02		0	
Split Info	0.371		0.863		0.94		0.98		0.94		0.863		0.651	
Gain ratio	0.129		0.017		0.047		0.091		0.108		0.023		0	

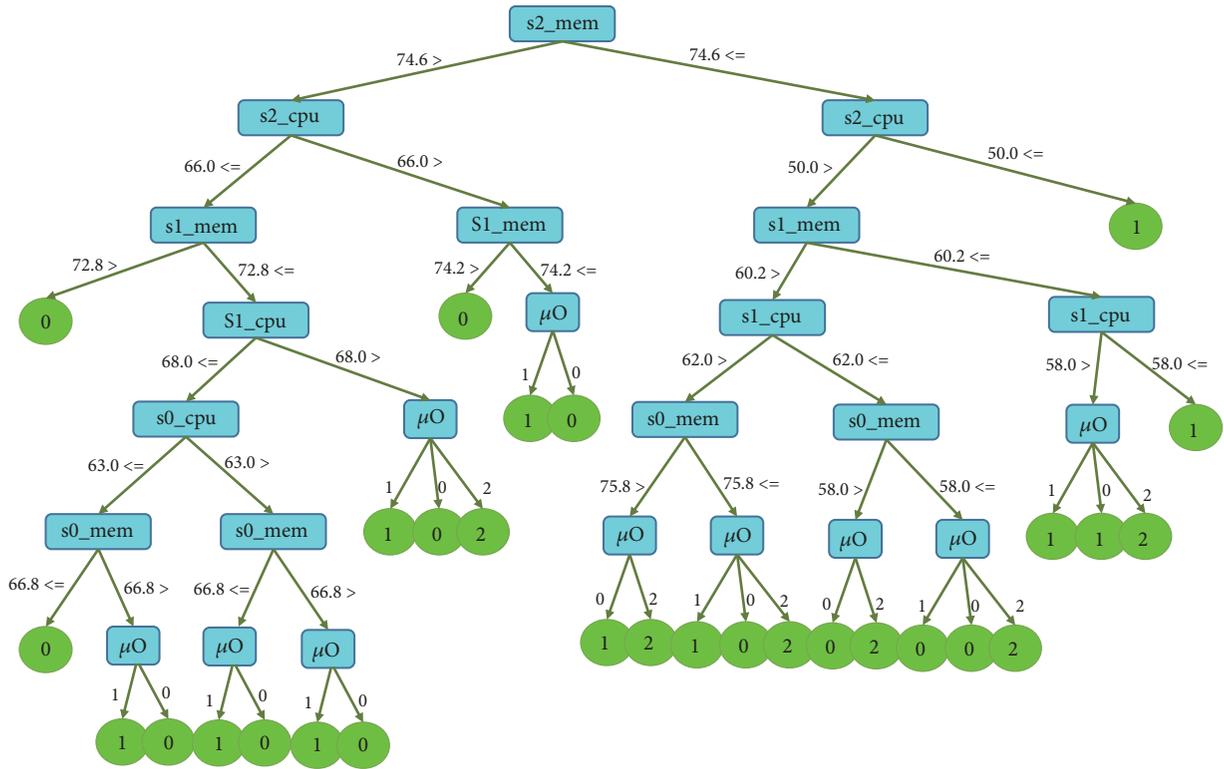


FIGURE 5: An example of established decision tree.

the subtree. The false positive number of leaf node is shown in (9).

$$E(\text{subtree\_error\_count}) = N * e \quad (7)$$

$$\text{var}(\text{subtree\_error\_count}) = \sqrt{N * e * (1 - e)} \quad (8)$$

$$E(\text{leaf\_error\_count}) = N * e \quad (9)$$

$$E(\text{sub\_error\_count}) - \text{var}(\text{subtree\_error\_count}) > E(\text{leaf\_error\_count}) \quad (10)$$

As shown in (10), if the number of misclassified subtrees minus the standard error of the error rate is still greater than the number of false positives corresponding to the leaf node; then the pruning will occur and the subtree will be removed from the tree. Until all nodes in all decision trees

have been visited, the pruning process is completed. The pruning process is finished when all the nodes in all the decision trees are visited. Figure 6 shows the pruned decision tree.

The proposed DTBFR mechanism utilizes the SDN technology to collect the network traffic information. The calculation results of decision tree algorithm can provide instructions for SDN controller to guide user traffic flows to the services in the nearby 5G  $\mu O$  micro data center. The DTBFR mechanism combined with 5G SDN/NFV programmable network architecture can improve service efficiency and save available bandwidth resources.

#### 4. Experiment Results

The purpose of this experiment is to verify  $\mu O$ 's communication infrastructure pattern during operation. The experiment

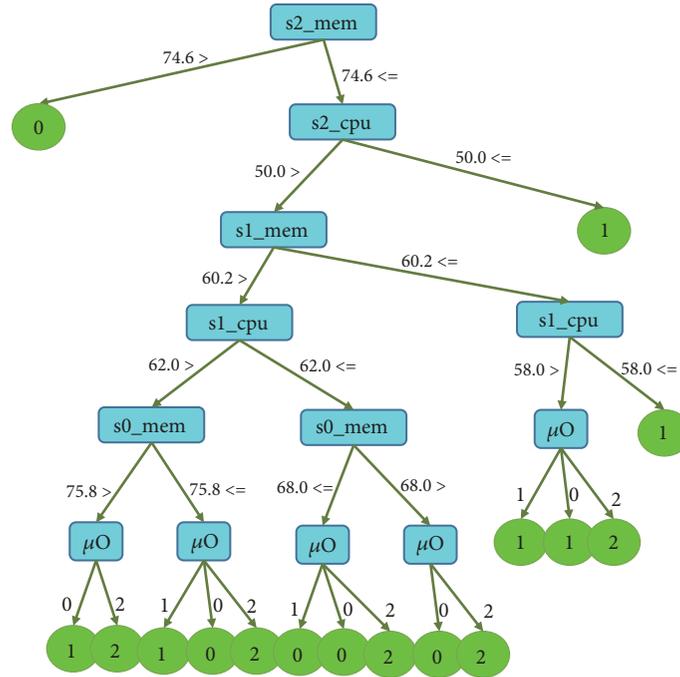


FIGURE 6: The pruned decision tree.

condition consists of multiple local area networks controlled by the  $\mu\text{O}$  and central tunnel controllers. In our evaluation, we simulated a network with tree-like topology in mininet and measured the overhead caused by network view division [42]. Mininet is used to simulate  $\mu\text{O}$ 's network, and tunnels can be built among different  $\mu\text{O}$  areas for quick exchange of information. The experiment comprises two hypervisors, one OpenVirtex, and multiple SDN controllers shown in Figure 7. The topology is constructed by the mininet in VM in the hypervisor and the  $\mu\text{O}$  areas are concatenated by tunneling technology. The OpenVirtex functions as the system management system. The IT staff can construct network slicing for the system via network interface and the back-end server of the network interface will analyse the operating information before interpreting it into OpenVirtex to produce commands for network slicing. The webpage uses socket and slice manager to transmit information. After the slice manager receives the information, it will continue to use OpenVirtex's API to build network slicing for the manager.

Network slicing has a wide range of applications according to the cohorts of users, the types of application services, or different demands for resources. The paper uses regionalization service as the benchmark for network slicing and connects homogeneous regionalization service servers with tunneling technology to form a complete network slicing of regionalization service. A C4.5 decision tree model is built based on the mechanism proposed by Ross Quinlan to serve as the reference for determination of 5G  $\mu\text{O}$ 's traffic flow direction and allows the users' traffic flow to be directed to a closer  $\mu\text{O}$  access service to reduce the network latency in packet transmission.

In order to verify  $\mu\text{O}$  0 and  $\mu\text{O}$  1's virtual communication service infrastructures, the experiment uses hypervisor to stimulate the scenario of adjacent  $\mu\text{O}$  communication. The infrastructure of the experiment consists of one Xen hypervisor and two Ubuntu 16.04 virtual machines. Mininet is used to virtualize S1 for VM1, add GRE interface, and set the remote IP as 10.10.10.126, which is the IP of VM2. VM2 has the same setting with the only difference that its remote IP is set as 10.10.10.116, which is the IP of VM1. Both VM1 and VM2's control plane are set as OpenVirtex and OVX will simultaneously finish logical network slicing when the slicing is built. As shown in Figure 8, the TCP bandwidth test is conducted by Iperf in mono-hypervisor. Iperf is a network benchmarking tool which is used to measure the throughput of the network carrying UDP and TCP data [43]. The average bandwidth of GRE tunnels is 1.65 Gbps and the average bandwidth of VxLAN tunnels is 1.63 Gbps. The results of the experiment are shown in Figure 9. As shown in Figure 10, logical network slicing infrastructure is built by the two hypervisors. Hypervisor 1 and Hypervisor 2 are separated physically. The results of the experiment are shown in Figure 11. In this case, the average bandwidth of GRE tunnels is 915 Mbps and the average bandwidth of VxLAN tunnels is 908 Mbps.

Judging from the results, we can see that even though tunnels between the two VMs in mono-hypervisor are successfully built, the packets are unable to actually leave hypervisor's network card but exchange memory instead. This is because both VMs are located in the same hypervisor. In the two hypervisors scenario, tunnels are successfully built as well and the packets actually reach each other's hypervisor network cards to carry out the exchange. This is attributed

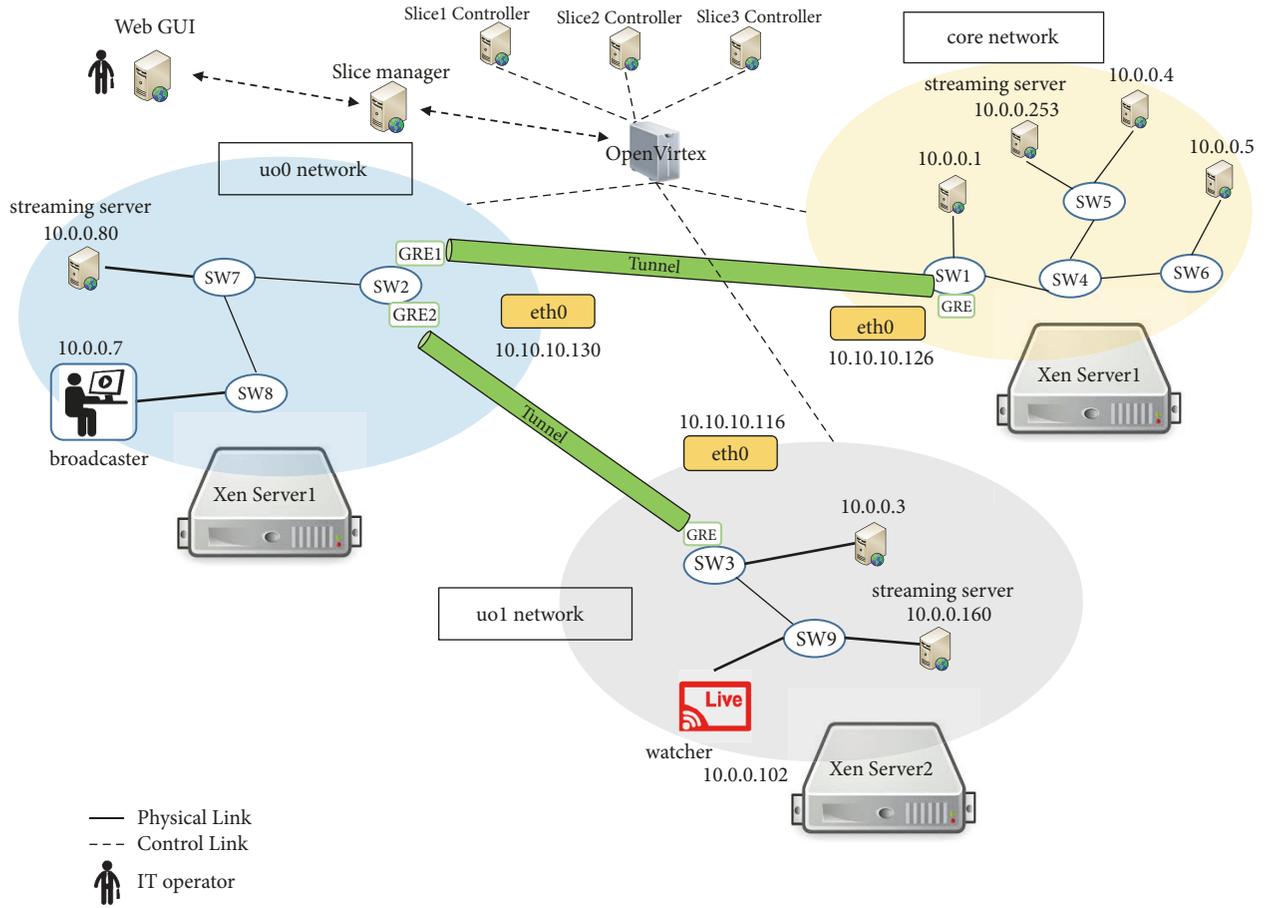


FIGURE 7: The experiment environment.

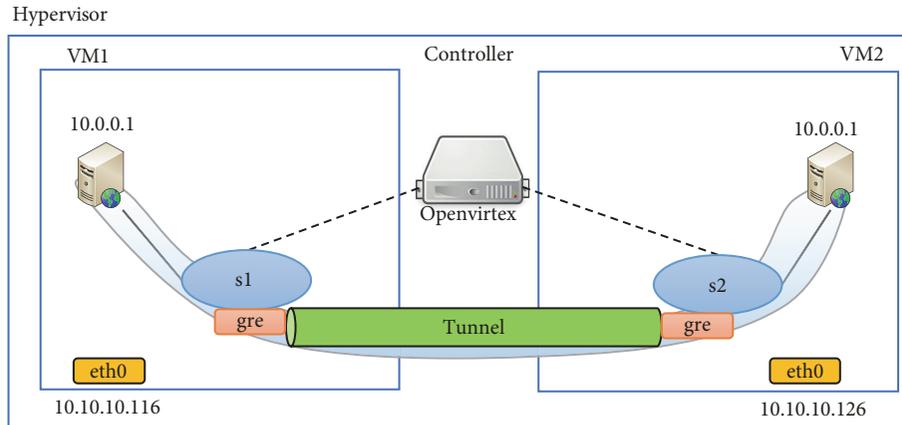


FIGURE 8: Average bandwidth test in a signal hypervisor scenario.

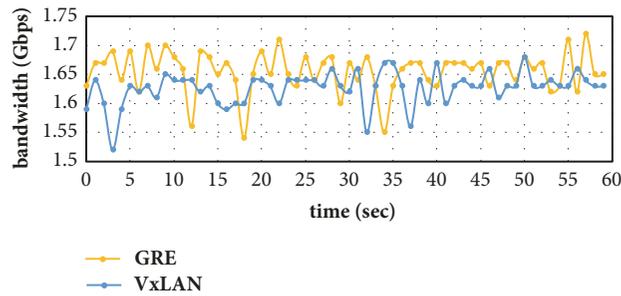


FIGURE 9: The experiment result in a signal hypervisor.

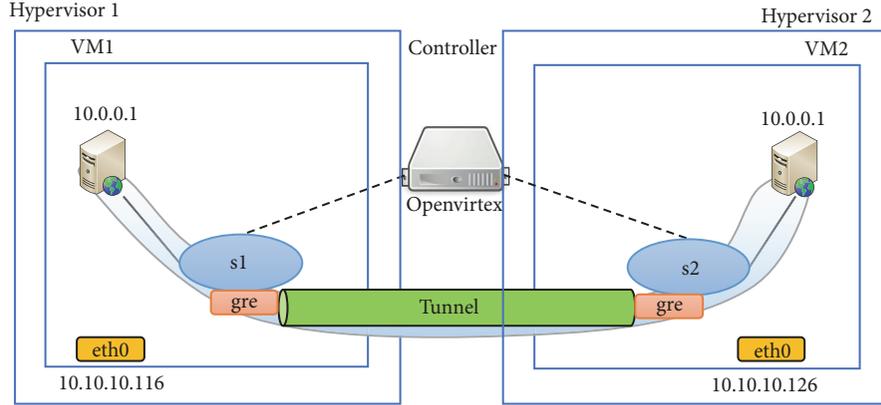


FIGURE 10: Average bandwidth test in two hypervisors scenario.

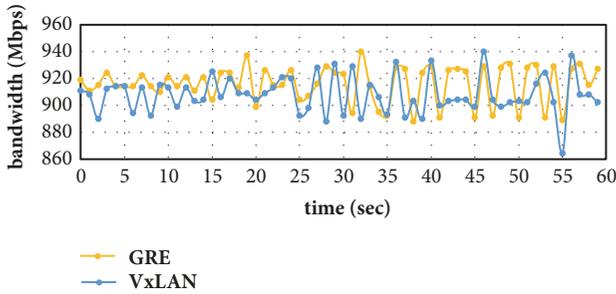


FIGURE 11: The experiment result in cross-hypervisors.

to that fact that a bridge of communication has been built by the tunnels and the packets are repackaged. It is therefore determined that the two hypervisors scenario requires less bandwidth than mono-hypervisor.

As for the scenario of nonadjacent  $\mu O$  communication, we replace the experiment infrastructure with two-tier mininet to simulate the whole environment and construct a virtual experiment infrastructure in different domains. As shown in Figure 12, in the first tier, we use the mininet to construct topology in different domains and connect domain 1 and domain 2 via the OpenVSwitch (OVS) of the mininet. Domain 1 and Domain 2 can also construct their own topology via the mininet. The topology of the mininet in the first tier uses OpenVirtex to conduct the network slicing for tenants. The slices are composed of tunnel and nontunnel, and the slices are subject to the management of different controllers. For example, the tunnel slice is subject to tunnel controller and so forth. The tunnel network slicing uses the exclusive network slices built by the tunneling technology. The experiment uses two mainstream tunneling technologies, GRE and VxLAN, to conduct the comparison. The experiment result is shown in Figure 13. The experiment sets the bandwidth of the link in the topology at 100Mbytes and the major purpose of the experiment is to test the transmission efficiency of the two tunneling technologies. GRE tunnel package belongs to the third tier routing; VxLAN uses UDP to pack the packet. The average bandwidth of GRE is around 73.7 Mbytes and the average bandwidth of

VxLAN is around 34.8 Mbytes. The results clearly suggest that VxLAN’s packet transmission efficiency is not even half of that of GRE’s. The cause behind this is because of the different ways of packet transmission. UDP is by itself an unreliable way of transmission and the packet is even more likely to be lost with the OpenVirtex serving as the intermediary tier. This might explain the poor performance in bandwidth.

To evaluate the efficiency and the performance, the DTBFR mechanism proposed in this paper is compared with the Minimize Loading First (MLF) mechanism. The MLF is a greedy algorithm that can be used to find the server with the lowest load and direct the traffic to the server via SDN controller.

In order to make users of  $\mu O$  gain access to the nearby micro data center or link to the Internet data center for access to particular application of network services via tunnels, three indicators are available to determine the quality of the decision-making of traffic flow redirection.  $\mu O\_num$  represents the number of accessed services in  $\mu O$ . Near- $\mu O\_num$  represents the number of times when the traffic flow has to access services in nearby  $\mu O$  because the servers in  $\mu O$  are all busy. Datacenter- $num$  represents the number of times recorded when the traffic flow has to access remote data center because nearby  $\mu O$ s are all busy. The results of the experiment are shown in Figure 14. The value of  $\mu O\_num$  of the proposed DTBFR mechanism is apparently higher than that of MLF mechanism. This means that DTBFR mechanism is able to prioritize the user’s traffic flow according to the servers in  $\mu O$ ’s area. If the server is busy, DTBFR mechanism will consider accessing services in nearby  $\mu O$  and, because the number of HTTP requests varies, (11) is used to calculate the redirection ratio.

$$\text{Ratio} = \frac{(\mu O\_num / \text{near-}\mu O\_num / \text{Datacenter\_num})}{\text{Total number of HTTP requests}} \quad (11)$$

Under the same experiment conditions, the decision-making of traffic flow direction of DTBFR mechanism is as follows: 68.1% of the traffic flow will be directed to the  $\mu O$  in the current area to access services; 28.7% of the traffic flow will be directed to the nearby  $\mu O$  to access services; only 3.2% of the traffic will go to the data center to access services.

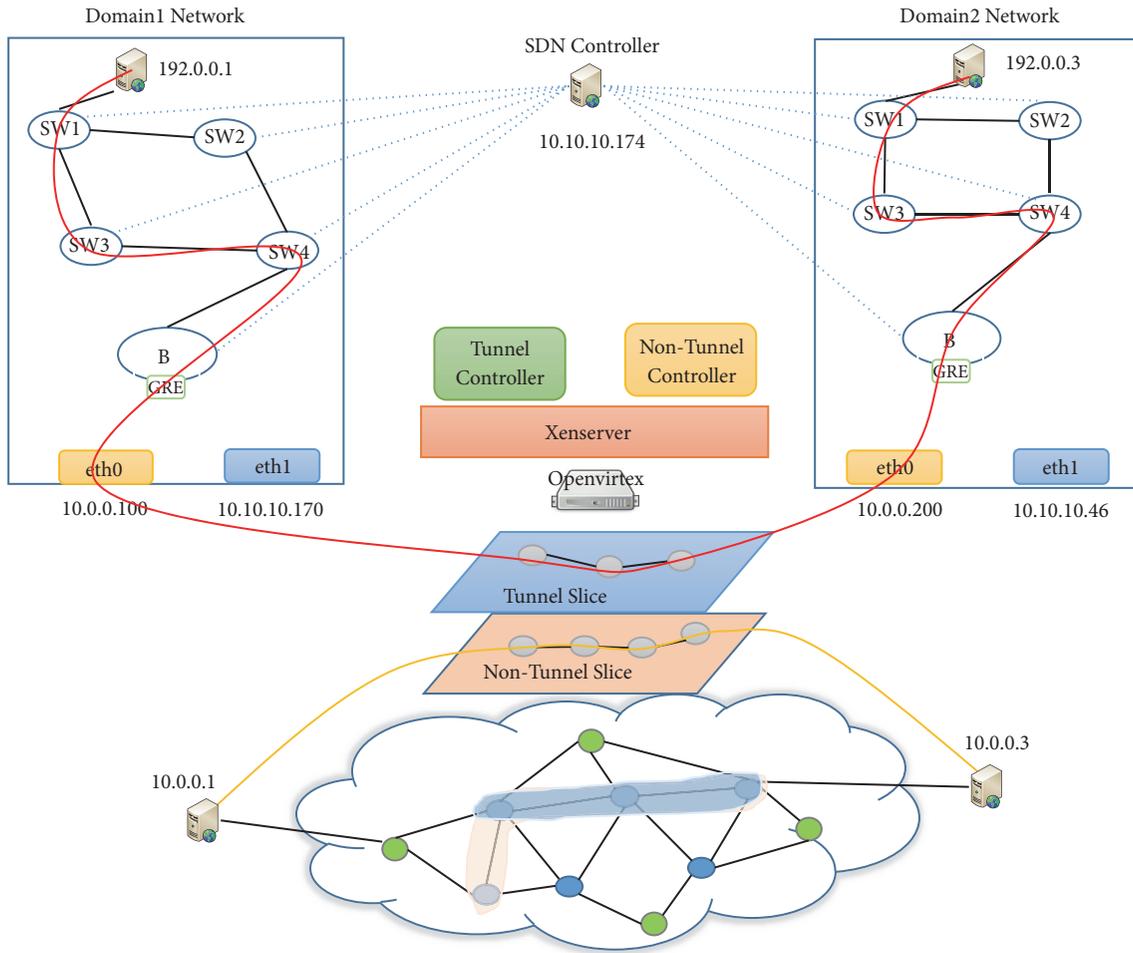


FIGURE 12: Nonadjacent  $\mu$ O communication scenario.

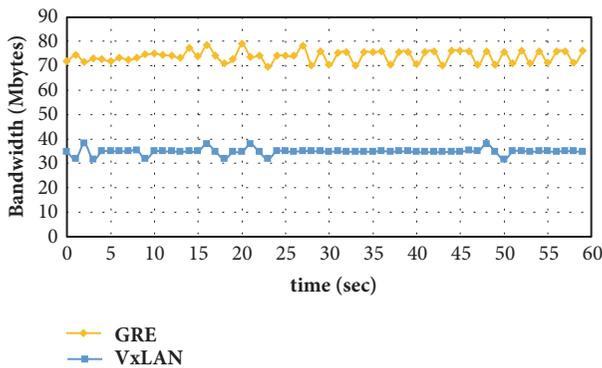


FIGURE 13: The comparison of GRE and VxLAN experiment result in nonadjacent  $\mu$ O communication scenario.

MLF algorithm is a directing method based on server load which directs 29.8% of the traffic flow to the  $\mu$ O in the current area to access services, 32.7% of the traffic flow to the nearby  $\mu$ O to access services, and 37.5% of the traffic flow to the Internet data center to access services. The  $\mu$ O's accessed services indicator in the area shows that DTBFR mechanism

takes up 68.1% while MLF algorithm takes up only 29.8%. The experiment verifies that DTBFR mechanism's performance on traffic redirection is relatively good.

In summary, the design pattern of 5G micro operator combines network slicing and tunneling technologies to realize the communication model for  $\mu$ O, although network slicing holds much promise for 5G, but not without its share of hurdles. The traditional network architecture will need to be redesigned to enable network slicing. Interoperability should be tested in order to ensure network slicing works as expected in the 5G network. In terms of tunneling, the inefficient packet overhead may have a negative effect on the performance of the network. For the purpose of micro operator to providing regional services, the DTBFR mechanism allows different  $\mu$ O's to share network resources and improves data processing by directing the users' traffic flow to a closer  $\mu$ O for reducing the network transmission latency, building the foundations of ultrareliable and low-latency communications (URLLC) in 5G. Conclusively, the  $\mu$ O design pattern provided in this paper helps to reduce the computing burden of the traditional cloud network architecture, improves the load capacity of the cloud network and the servers, and is also able to be used as development basis for

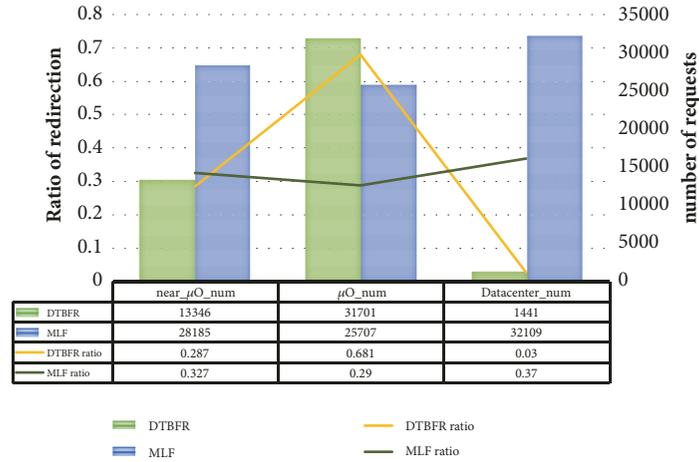


FIGURE 14: Comparison with experiment results of DTBFR and MLF.

integration and innovation application of the Cloudy-Edge Computing achieved by 5G.

## 5. Conclusions

This paper uses SDN and NFV technologies as the base and combines network slicing and tunneling technologies to come up with a network infrastructure pattern for  $\mu$ O. The pattern allows users of different  $\mu$ O to be concatenated via tunneling technology and then realizes the rapid connection of network to effectively enhance the interconnectivity of networks. In order to meet the demand of regional service of micro operator, this paper proposes DTBFR mechanism, which uses decision tree theory as the basis of SDN-based traffic decision-making. Under the same experiment conditions, the decision-making of traffic flow direction of DTBFR mechanism is as follows: 68.1% of the traffic flow will be directed to the  $\mu$ O in the current area to access services; 28.7% of the traffic flow will be directed to the nearby  $\mu$ O to access services; only 3.2% of the traffic will go to the Internet cloud data center to access services. The features deployed by the regional service of the micro operator can effectively reduce the burden of the data center on the Internet and accelerate the development of the edge computing service in the future 5G network.

Future work is needed to create an edge computing service model for the wide-spread adoption of the  $\mu$ O design pattern, especially the integration of the edge computing technologies into the 5G  $\mu$ O deployment path by improving management efficiency, and provides service on-demand application for micro operator customers. The resource estimation, tasking scheduling, and lightweight virtual network function configuration techniques are the primary research directions.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The work described in this paper was supported in part by the Ministry of Science and Technology of the Republic of China (Project nos. 104-2221-E-008 -039, 105-2221-E-008-071, 107-2623-E-008-002, and 107-2636-E-003-001).

## References

- [1] X. Huang, Y. Li, S. Tang, and Q. Chen, "Coexistence of cognitive small cell and WiFi system: a traffic balancing dual-access resource allocation scheme," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 4092681, 17 pages, 2018.
- [2] A. Raschella, F. Bouhafs, G. C. Deepak, and M. Mackay, "QoS aware radio access technology selection framework in heterogeneous networks using SDN," *Journal of Communications and Networks*, vol. 19, no. 6, pp. 577–586, 2017.
- [3] H. Beyranvand, M. Levesque, M. Maier, J. A. Salehi, C. Verikoukis, and D. Tipper, "Toward 5G: FiWi enhanced LTE-A hetnets with reliable low-latency fiber backhaul sharing and WiFi offloading," *IEEE/ACM Transactions on Networking*, vol. 25, no. 2, pp. 690–707, 2017.
- [4] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "NFV and SDN-Key technology enablers for 5G networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2468–2478, 2017.
- [5] C. W. Tseng, P. H. Lai, B. S. Huang, L. D. Chou, and M. C. Wu, "NFV deployment strategies in SDN network," *International Journal of High Performance Computing and Networking (IJHPCN)*, 2018, Available: <http://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=ijhpcn>.
- [6] E. J. Kitindi, S. Fu, Y. Jia, A. Kabir, and Y. Wang, "Wireless network virtualization with SDN and C-RAN for 5G networks: requirements, opportunities, and challenges," *IEEE Access*, vol. 5, pp. 19099–19115, 2017.

- [7] uO5G, “Micro operator concept for boosting local service delivery in 5G,” Available: <http://www oulu.fi/uo5g/>.
- [8] M. Matinmikko, M. Latva-aho, P. Ahokangas, S. Yrjölä, and T. Koivumäki, “Micro operators to boost local service delivery in 5G,” *Wireless Personal Communications*, vol. 95, no. 1, pp. 69–82, 2017.
- [9] A. Prasad, Z. Li, S. Holtmanns, and M. A. Uusitalo, “5G micro-operator networks — A key enabler for new verticals and markets,” in *Proceedings of the 2017 25th Telecommunication Forum (TELFOR)*, pp. 1–4, Belgrade, Serbia, November 2017.
- [10] J. S. Walia, H. Hammainen, and M. Matinmikko, “5G Micro-operators for the future campus: A techno-economic study,” in *Proceedings of the 2017 Internet of Things - Business Models, Users, and Networks*, pp. 1–8, Copenhagen, Denmark, November 2017.
- [11] M. Matinmikko-Blue and M. Latva-aho, “Micro operators accelerating 5G deployment,” in *Proceedings of the 2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, pp. 1–5, Peradeniya, Sri Lanka, December 2017.
- [12] K. B. Manosha, M. Matinmikko-Blue, and M. Latva-aho, “Framework for spectrum authorization elements and its application to 5G micro-operators,” in *Proceedings of the 2017 Internet of Things - Business Models, Users, and Networks*, pp. 1–8, Copenhagen, Denmark, November 2017.
- [13] T. Sanguanpuak, S. Guruacharya, E. Hossain, N. Rajatheva, and M. Latva-aho, “On spectrum sharing among micro-operators in 5G,” in *Proceedings of the 2017 European Conference on Networks and Communications (EuCNC)*, pp. 1–6, Oulu, Finland, June 2017.
- [14] P. Ahokangas, S. Moqaddamerad, and M. Matinmikko, “Future micro operators business models in 5G,” *The Business and Management Review*, vol. 7, no. 5, pp. 143–149, 2016.
- [15] P. Mach and Z. Becvar, “Mobile Edge Computing: A Survey on Architecture and Computation Offloading,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.
- [16] G. Li, J. Song, J. Wu, and J. Wang, “Method of resource estimation based on QoS in edge computing,” *Wireless Communications and Mobile Computing*, vol. 2018, 9 pages, 2018.
- [17] M. A. Lema, A. Laya, T. Mahmoodi et al., “Business case and technology analysis for 5G low latency applications,” *IEEE Access*, pp. 5917–5935, 2017.
- [18] P. Kiss, A. Reale, C. J. Ferrari, and Z. Istenes, “Deployment of IoT applications on 5G edge,” in *Proceedings of the 2018 IEEE International Conference on Future IoT Technologies (Future IoT)*, pp. 1–9, Eger, Hungary, January 2018.
- [19] M. Matinmikko, A. Roivainen, M. Latva-aho, and K. Hiltunen, “Interference study of micro licensing for 5g micro operator small cell deployments,” in *Cognitive Radio Oriented Wireless Networks*, vol. 228 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 264–275, Springer International Publishing, Switzerland, 2018.
- [20] L. Falconetti, R. Karaki, and S. Corroy, “Practical energy-aware cell association for small cell deployment,” *Wireless Communications and Mobile Computing*, vol. 16, no. 16, pp. 2436–2448, 2016.
- [21] T. Zahid, X. Hei, We. Cheng, A. Ahmad, and P. Maruf, “On the tradeoff between performance and programmability for software defined WiFi networks,” *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 1083575, pp. 1–12, 2018.
- [22] P. Ahokangas, M. Matinmikko, S. Yrjölä, and I. Atkova, “Disruptive revenue models for future micro operator driven mobile business ecosystem,” in *Proceedings of The 24th Nordic Academy of Management Conference (NFF)*, pp. 23–25, Bodo, Norway, 2017.
- [23] M. G. Kibria, G. P. Villardi, K. Nguyen, W.-S. Liao, K. Ishizu, and F. Kojima, “Shared spectrum access communications: a neutral host micro operator approach,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1741–1753, 2017.
- [24] “Micro operators for vertical specific service delivery in 5G,” Available: <http://www.5gsummit.org/berlin/docs/slides/Matti-Latva-Aho.pdf/>.
- [25] A. Basta, A. Blenk, K. Hoffmann, H. J. Morper, M. Hoffmann, and W. Kellerer, “Towards a cost optimal design for a 5G mobile core network based on SDN and NFV,” *IEEE Transactions on Network and Service Management*, vol. 14, no. 4, pp. 1061–1075, 2017.
- [26] C. Bouras, A. Kollia, and A. Papazois, “SDN and NFV in 5G: Advancements and challenges,” in *Proceedings of the 20th Conference on Innovations in Clouds, Internet and Networks, ICIN 2017*, pp. 107–111, France, March 2017.
- [27] ONF, “SDN architecture,” Available: <https://www.opennetworking.org/images/stories/downloads/sdnresources/technical-reports/TRSDNARCH1.006062014.pdf>.
- [28] ETSI, “Network Functions Virtualization (NFV), architectural framework,” ETSI Ind. Specification Group, Sophia Antipolis Cedex, France, 2014.
- [29] Q. Jia, R. Xie, T. Huang, J. Liu, and Y. Liu, “Efficient caching resource allocation for network slicing in 5G core network,” *IET Communications*, vol. 11, no. 18, pp. 2792–2799, 2017.
- [30] X. Li, M. Samaka, H. A. Chan et al., “Network slicing for 5G: challenges and opportunities,” *IEEE Internet Computing*, vol. 21, no. 5, pp. 20–27, 2017.
- [31] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, “From network sharing to multi-tenancy: The 5G network slice broker,” *IEEE Communications Magazine*, vol. 54, no. 7, pp. 32–39, 2016.
- [32] NGMN Alliance, “5G White Paper,” Available: <https://www.ngmn.org/5g-white-paper.html>.
- [33] NOKIA, “Dynamic end-to-end network slicing for 5G White Paper,” Available: [http://www.hit.bme.hu/~jakab/edu/litr/5G/NOKIA\\_dynamic\\_network\\_slicing-WP.pdf](http://www.hit.bme.hu/~jakab/edu/litr/5G/NOKIA_dynamic_network_slicing-WP.pdf).
- [34] S. Min, S. Kim, J. Lee, B. Kim, W. Hong, and J. Kong, “Implementation of an OpenFlow network virtualization for multi-controller environment,” in *Proceedings of the 14th International Conference on Advanced Communication Technology, ICACT 2012*, pp. 589–592, Republic of Korea, February 2012.
- [35] A. Al-Shabibi, M. De Leenheer, M. Gerola et al., “OpenVirteX: Make your virtual SDNs programmable,” in *Proceedings of the 3rd ACM SIGCOMM 2014 Workshop on Hot Topics in Software Defined Networking, HotSDN 2014*, pp. 25–30, Chicago, IL, USA, August 2014.
- [36] E. Alpaydin, *Introduction to Machine Learning*, The MIT Press, Cambridge, Mass, London, England, 2nd edition, 2009.
- [37] B. Hssina, A. Merbouha, H. Ezzikouri, and M. Erritali, “A comparative study of decision tree ID3 and C4.5,” *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 2, 2014, Special Issue on Advances in Vehicular Ad Hoc Networking and Applications 2014.
- [38] R. C. Barros, M. P. Basgalupp, A. C. P. L. F. De Carvalho, and A. A. Freitas, “A survey of evolutionary algorithms for decision-tree induction,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 42, no. 3, pp. 291–312, 2012.

- [39] D. Lee and C. S. Hong, "Access point selection algorithm for providing optimal AP in SDN-based wireless network," in *Proceedings of the 2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 362–365, Seoul, South Korea, September 2017.
- [40] S. T. V. Pasca, S. S. P. Kodali, and K. Kataoka, "AMPS: Application aware multipath flow routing using machine learning in SDN," in *Proceedings of the Twenty-third National Conference on Communications (NCC)*, pp. 1–6, Chennai, India, 2017.
- [41] "Entropy," Available: [https://en.wikipedia.org/wiki/Entropy\\_\(information\\_theory\)](https://en.wikipedia.org/wiki/Entropy_(information_theory)).
- [42] "Mininet," Available: <http://mininet.org/>.
- [43] "iPerf," Available: <https://iperf.fr/>.

