IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 5, 2021, accepted February 19, 2021, date of publication March 2, 2021, date of current version April 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3063177

Reliable Virtual Network Mapping Algorithm With Network Characteristics and Associations

SHUNLI ZHANG
Department of Information Technology and Engineering, Jinzhong University, Jinzhong 030619, China
State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecom, Beijing 100876, China
e-mail: 544891117@qq.com

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1800704.

ABSTRACT In the existing research on reliable virtual network mapping, the relationship between network parameters is insufficiently analyzed, the network topology and historical mapping data are not fully utilized, resulting in a low virtual network mapping success rate and low utilization of the underlying network resources. To resolve this problem, this article proposes a reliable virtual network mapping algorithm based on network features and associations. Firstly, the network features related to reliable virtual network mapping are sorted out, and the underlying node reliability matrix is established based on historical data. Secondly, a Bayesian network-based inference model is constructed. Finally, based on the underlying node reliability matrix and virtual network features, two algorithms named NFA-TS and NFA-LR about reliable virtual network mapping were proposed. Compared with two algorithms named SVNE and VNE-SSM, the proposed algorithm shows the competitive performance on virtual network mapping success rate and the underlying network resource utilization.

INDEX TERMS Virtual network, underlying network, virtual network mapping, mapping algorithms, network characteristics, associations.

I. INTRODUCTION

In recent years, network virtualization has become a key technology to solve the problem of network rigidity [1], [2]. In a network virtualization environment, traditional physical networks are divided into underlying networks and virtual networks. Multiple heterogeneous virtual networks can be mapped to one underlying network at the same time, providing rich and diverse services to end users. Virtual network resource allocation has become a key research hotspot in network virtualization.

From the perspective of network reliability analysis, current virtual network mapping algorithms can be divided into two types: mapping algorithms that do not consider reliability [3]–[9], and mapping algorithms that consider reliability [10]–[15]. Through analysis, it can be seen that the mapping algorithm considering reliability is more in line with the actual situation of the network and has good application value.

However, the existing research that consider reliability has the following two shortcomings: (1) In order to select the underlying network resources with high reliability, the existing research uses the strategy of adding the characteristics parameters of the underlying network resources to find the underlying network with high reliability. Allocating resources, ignoring the relationship between the network characteristics of the virtual network and the success rate of virtual network mapping, cannot effectively measure the importance of some parameters. Every time by selecting the underlying resource with the largest resource for allocation, it is easy to make the underlying network lack Large-capacity underlying resources, which causes the virtual network request mapping that requires a large number of resources to fail. (2) Existing studies have selected highly reliable underlying network resources as backup links or redundant links. This resource allocation strategy lacks an effective analysis of existing virtual network mapping data, resulting in more critical underlying network resources being used as backup resources, affecting the success rate of virtual network mapping and the utilization of the underlying network resources.

In order to solve these problems, this article proposes a reliable virtual network mapping algorithm based on network characteristics and association relationships. The innovations include:

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Wang.

(1) Sorting out the characteristics of the virtual network and underlying network related to reliable virtual network mapping.

(2) Based on the historical data of virtual network mapping, the association relationship between network features is mined, and the underlying node reliability matrix based on historical information is established.

(3) Establish an inference model based on the Bayesian network, which is used to solve the optimized underlying node that the current virtual node's neighbor nodes can map on the premise that the underlying node that the current virtual node maps is known.

(4) In order to make full use of the reliability matrix of the underlying nodes, a two-stage mapping algorithm NFA-TS that preferentially maps virtual nodes and a hierarchical network-based mapping algorithm NFA-LR are proposed, which are compared with existing research results. The effectiveness of the proposed algorithm in terms of the success rate of virtual network request mapping and the utilization of the underlying network resources is verified.

II. RELATED WORK

Virtual network mapping is a key issue in network virtualization research. From the perspective of network reliability, the current virtual network mapping algorithms can be divided into two types: mapping algorithms that do not consider reliability, and mapping algorithms that consider reliability.

In terms of mapping algorithms that do not consider reliability, the researchers' main goal is to improve the utilization of the underlying network resources, which can be divided into two types to improve resource utilization and reduce energy consumption. (1) In terms of improving resource utilization: Literature [3] reduces the occupation of the underlying link bandwidth resources by the virtual link by establishing a mathematical model, and generates a resource allocation strategy that maximizes the utilization of basic network resources. Literature [4] proposed a heuristic optimization algorithm for optimal allocation of basic network resources. Literature [5] uses an improved genetic algorithm to solve the virtual network mapping problem. Reference [6] allows the virtual network to dynamically change its resource requirements during resource allocation, thereby maximizing the utilization of the underlying network resources. Reference [7] applied deep learning algorithms to the study of virtual network resource allocation. By encoding physical networks and virtual networks into images, the problem characteristics of automatic selection of deep reinforcement learning methods were realized, and the application range of the algorithm was increased. (2) Reduction of energy consumption: Literature [8] proposed a self-organizing mapping neural network algorithm, which saved the energy consumption of the underlying network. Literature [9] modeled the virtual network mapping problem as a multi-objective optimization problem, and solved it through an improved genetic algorithm, which reduced the energy consumption of the underlying network better.

In terms of the reliability of the mapping algorithm, the researchers' main goal is to improve the utilization of the underlying network resources while ensuring the reliability of the virtual network. Literature [10], [11] summed up the parameter values affecting the reliability of the underlying network to obtain a highly reliable underlying network resource, and proposed a reliability-aware virtual network mapping algorithm, which achieved good results in the efficiency and reliability of virtual network mapping. Literature [12], [13] proposed a virtual network mapping algorithm with anti-destruct capability, which maps the virtual network to the underlying network resources with strong anti-destruct capability. Literature [14] takes energy consumption and availability as the goal of resource allocation, and proposes a greedy stochastic adaptive search algorithm to ensure the quality-of-service requirements of each virtual network. Literature [15] first divided the underlying network resources into two types: available resources and backup resources. Secondly, a heuristic mapping algorithm is proposed to allocate available resources to virtual network requests. When the underlying network fails, backup resources are used instead of the failed resources. This algorithm improves the reliability of the virtual network.

Through analysis, it can be seen that the mapping algorithm considering reliability is more in line with the actual situation of the network and has good application value. However, in existing mapping algorithms that consider reliability, the relationship between network parameters is insufficiently analyzed, network topology and historical mapping data are not fully utilized, resulting in a low success rate of virtual network mapping and low utilization of the underlying network resources. In order to solve this problem, this article sorts out the network characteristics related to reliable virtual network mapping, establishes the underlying node reliability matrix and inference model based on historical data, and proposes two-stage mapping algorithm NFA-TS that preferentially maps virtual nodes, and virtual network mapping algorithm based on hierarchical relationships NFA-LR.

III. PROBLEM DESCRIPTION

A. NETWORK DESCRIPTION

1) UNDERLYING NETWORK

The weighted undirected graph $G_S = (N_S, E_S)$ is used to represent the underlying network. Among them, N_S represents the underlying node set, and each underlying node $n_i^s \in N_S$ contains attributes: CPU resource $cpu(n_i^s)$, location $loc(n_i^s)$. E_S represents a set of underlying links, and the attribute contained in each underlying link $e_j^s \in E_S$ is a bandwidth resource $bw(e_j^s)$. Let $|N_S|$ be the number of underlying nodes. Use $|E_S|$ to denote the number of underlying links.

2) VIRTUAL NETWORK REQUEST

Use weighted undirected graph $G_V = (N_V, E_V)$ to represent virtual network requests. Among them, N_V represents a set of virtual nodes, and the attribute of each virtual node

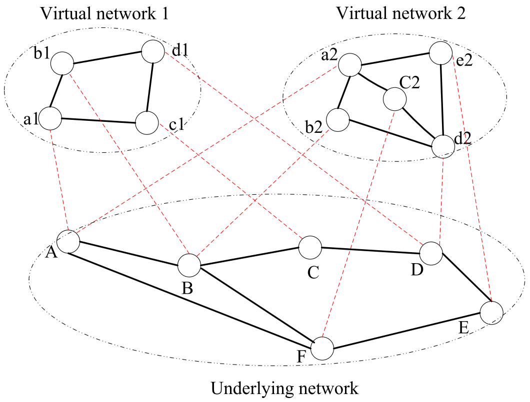


FIGURE 1. Virtual network mapping.

$n_i^v \in N_V$ is CPU resource $cpu(n_i^v)$. E_V represents a set of virtual links. The attribute of each virtual link $e_j^v \in E_V$ is bandwidth resource $bw(e_j^v)$. Let $|N_V|$ be the number of virtual nodes and $|E_V|$ be the number of virtual links.

3) VIRTUAL NETWORK MAPPING

Virtual network mapping means that the underlying network service provider selects resources from the underlying network that meet the virtual network request conditions and allocates them to the virtual network according to the resource requirements requested by the virtual network. Virtual network mapping includes virtual node mapping and virtual link mapping. In the virtual node mapping phase, the CPU resource $cpu(n_i^s)$ of the underlying node $n_i^s \in N_S$ allocated to the virtual node $n_i^v \in N_V$ needs to meet the CPU resource requirement $cpu(n_i^v)$ of the virtual node. Use $n_i^v \downarrow n_i^s$ to indicate that the virtual node $n_i^v \in N_V$ is mapped to the underlying node $n_i^s \in N_S$.

In the virtual link mapping phase, the bandwidth resources $bw(e_j^s)$ of all the underlying links $e_j^s \in E_S$ allocated to the virtual link $e_j^v \in E_V$ need to meet the bandwidth resource requirements $bw(e_j^v)$ of the virtual link. Use $e_j^v \downarrow p_j^s$ to indicate that virtual link $e_j^v \in E_V$ is mapped to the underlying path $p_j^s \in E_S$. Among them, p_j^s represents a set of underlying links composed of the physical link $e_j^s \in E_S$ allocated to the virtual link $e_j^v \in E_V$. **Fig. 1**. contains an underlying network and two virtual networks. Take virtual network 1 as an example to explain the virtual network mapping process: (1), node mapping: $\{a1 \rightarrow A, b1 \rightarrow B, c1 \rightarrow C, d1 \rightarrow D\}$; (2), link mapping:

$$\{(a1, b1) \rightarrow (A, B); (a1, c1) \rightarrow (A, B), (B, C);$$
$$(b1, d1) \rightarrow (B, C), (C, D); (c1, d1) \rightarrow (C, D)\}.$$

B. EVALUATION INDEX

1) UNDERLYING NETWORK COST

The cost of the underlying network refers to the sum of the node resources and link resources of the underlying network occupied by the successfully mapped virtual network at a time t. Let C_t^S be the cost of the underlying network at time t, and use formula (1) to calculate it. Among them, $hop(e_j^v)$ represents the number of underlying links to which e_j^v is mapped. Use C_T^S to represent the cost of the underlying network during time period T, and use formula (2) to calculate it.

$$C_t^S = \sum\nolimits_{n_i^v \in N_V} cpu(n_i^v) + \sum\nolimits_{e_i^v \in E_V} hop(e_j^v) \times bw(e_j^v) \quad (1)$$

$$C_T^S = \lim_{T \to \infty} \frac{\sum_{t=0}^{T} C_t^S}{T} \quad (2)$$

2) UNDERLYING NETWORK PROFIT

The underlying network profit refers to the sum of virtual node resources and virtual link resources that are successfully mapped at time t. Use R_t^S represents underlying network revenue a time t, and use formula (3) to calculate it. Use R_T^S represents underlying network revenue during the time period T, and use formula (4) to calculate it.

$$R_t^S = \sum\nolimits_{n_i^v \in N_V} cpu(n_i^v) + \sum\nolimits_{e_j^v \in E_V} bw(e_j^v) \quad (3)$$

$$R_T^S = \lim_{T \to \infty} \frac{\sum_{t=0}^{T} R_t^S}{T} \quad (4)$$

3) RATIO OF REVENUE TO EXPENSE

The ratio of revenue to expense of the underlying network resources refers to the revenue of the underlying network divided by the cost of the underlying network during the time period T, which is represented by U_T^S and calculated using formula (5).

$$U_T^S = \frac{R_T^S}{C_T^S} \quad (5)$$

4) MAPPING SUCCESS RATE

The mapping success rate refers to the number of virtual networks successfully mapped divided by the total number of virtual networks during the time period T, which is represented by Q_{win}^V and calculated using formula (6). Among them, $Q_{win}^V(t)$ represents the number of virtual network requests successfully mapped at time t; $Q^V(t)$ represents the total number of virtual network requests at time t.

$$Q_{win}^V = \lim_{T \to \infty} \frac{\sum\limits_{t=0}^{T} Q_{win}^V(t)}{\sum\limits_{t=0}^{T} Q^V(t)} \quad (6)$$

IV. NETWORK CHARACTERISTICS ANALYSIS

A. BASIC CHARACTERISTICS OF THE NETWORK

Because both the underlying node and the virtual node include the characteristics of the bandwidth resources of the adjacent links, for convenience of description, n_i is used to represent the underlying node and the virtual node. Let $AL(n_i)$ be the bandwidth resource of the adjacent link of node n_i and be calculated using formula (7). Among them, $E(n_i)$ represents the set of adjacent links of node n_i. The richer the bandwidth resources of adjacent links of node, the greater the success rate of virtual link mapping.

$$AL(n_i) = \sum\nolimits_{e_j \in E(n_i)} bw(e_j) \tag{7}$$

B. CHARACTERISTICS OF VIRTUAL NETWORKS

1) CENTER VALUE OF THE VIRTUAL NODE

Use $NC(n_i^v)$ to represent the central value of the virtual node. The calculation formula is shown in (8), where $hops(n_i, n_j)$ represents the number of hops from virtual node n_i^v to virtual node n_j^v. It can be known from formula (8) that $NC(n_i^v)$ represents the reciprocal of the sum of the number of hops of virtual node n_i^v to other virtual nodes n_j^v of the virtual network. The larger the value, the closer the current virtual node is to all virtual nodes, and the more likely it is to become the central node of the virtual network.

$$NC(n_i^v) = \frac{1}{\sum_{n_j^v \in N_V} hops(n_i^v, n_j^v)} \tag{8}$$

2) IMPORTANCE OF VIRTUAL NODE

Use $IMPORT(n_i^v)$ to represent the importance of virtual node $n_i^v \in N_V$, and use formula (9) to calculate. The virtual node with a large resource requirement and a large center value has a greater importance in the virtual network.

$$IMPORT(n_i^v) = (cpu(n_i^v) + AL(n_i^v)) \times NC(n_i^v) \tag{9}$$

C. CHARACTERISTICS OF THE UNDERLYING NETWORK

1) NUMBER OF HOPS FROM THE CANDIDATE UNDERLYING NODE TO THE MAPPED NODE

Let $UH(n_i^s)$ be the number of hops from the candidate underlying node n_i^s of the virtual node to the underlying node to which the virtual network has been mapped, and be calculated using formula (10). Among them, $\mu(n_i^s)$ represents the underlying node set that has been mapped successfully when the current underlying node n_i^s is selected as a candidate node. $hop(n_i^s, n_j^s)$ represents the number of hops from the candidate underlying node n_i^s of the currently-mapped virtual node to the underlying node n_j^s of the mapped virtual node. The smaller the number of hops from a node to a mapped neighbor node, the fewer the number of underlying links occupied by the virtual link mapping, and the higher the utilization of the underlying resources.

$$UH(n_i^s) = \sum\nolimits_{n_j^s \in \mu(n_i^s)} hop(n_i^s, n_j^s) \tag{10}$$

2) FAILURE RATE OF THE UNDERLYING NODES

Let $FN(n_i^s)$ be the failure rate of the underlying node $n_i^s \in N_S$. The higher the node failure rate, the more failures occur, indicating that the reliability of the node is lower. The failure rate of nodes can be solved by the operation and maintenance data of the network.

3) RESOURCE UTILIZATION OF THE UNDERLYING NODES

Use $RU(n_i^s)$ to indicate the resource utilization of the underlying node $n_i^s \in N_S$, which refers to the ratio of the CPU resources and adjacent link bandwidth resources of the used nodes to the total node CPU resources and adjacent link bandwidth resources. The higher the node's resource utilization, the lower the reliability of the node, and the more serious of the failure consequences.

4) RELIABILITY OF THE UNDERLYING NODES

Use $RELIAB(n_i^s)$ to represent the reliability of the underlying node $n_i^s \in N_S$, and calculate it using formula (11). Among them, the first half of the formula represents the reliability of the node's resource quantity, and the second half of the formula represents the reliability of the node's performance. κ and λ are tuning parameters.

$$RELIAB(n_i^s) = \kappa \frac{cpu(n_i^s) + AL(n_i^s)}{UH(n_i^s)} \times \lambda \frac{1}{FN(n_i^s) \times RU(n_i^s)} \tag{11}$$

V. UNDERLYING NODE RELIABILITY MATRIX

In the long-term operation of the underlying network service provider, a large amount of underlying network information and virtual network mapping information has been accumulated. Among them, the underlying network information includes the reliability and utilization of the underlying network resources. The virtual network mapping information includes information such as the arrival of the virtual network request, the end of the life cycle of the virtual network, and the mapping success rate. Based on these data, this section uses machine learning algorithms for data mining to establish the underlying node reliability matrix [16], [17].

1. Node CPU resource mapping experience matrix M_{CPU}

Use an n * n matrix M_{CPU} to indicate the importance of the CPU resources of the underlying nodes. The element values of the matrix are represented by $a_{ii} \in M_{CPU}$. The value of a_{ii} represents the sum of the CPU resources allocated by the underlying node $n_i^s \in N_S$ to all virtual nodes during the period T. The larger the value of a_{ii}, the more likely the new virtual node is mapped to the current underlying node, so the more important the current underlying node is. The value of $a_{ij} \in M_{CPU}$ is zero.

2. Node link resource mapping experience matrix M_{LINK}

An n * n matrix M_{LINK} is used to represent the importance of link resources of the underlying nodes. The element values of the matrix are represented by $b_{ij} \in M_{LINK}$. The value of b_{ij} represents the number of bandwidth resources allocated by path $P(n_i^s, n_j^s)$ to virtual link $e_k^v \in E_V$ during time period T divided by the number of hops of path $P(n_i^s, n_j^s)$. Among them, $P(n_i^s, n_j^s)$ represents a path from the underlying node $n_i^s \in N_S$ to the underlying node $n_j^s \in N_S$. The virtual nodes at both ends of the virtual link $e_k^v \in E_V$ are mapped to the underlying node $n_i^s \in N_S$ and the underlying node $n_j^s \in N_S$, respectively.

3. Node reliability experience matrix M_{RELIAB}

An n * n matrix M_{RELIAB} is used to represent the importance of the reliability of the underlying nodes. The element values of the matrix are represented by $c_{ii} \in M_{RELIAB}$. The value of c_{ii} represents the node reliability of the underlying node $n_i^s \in N_S$ in the time period T, and is calculated using formula (11). The larger the value of c_{ii}, the higher the reliability of the underlying node, and the more likely the new virtual node will be mapped to the current underlying node. Therefore, the more important the current underlying node is.

In order to establish the reliability matrix of a node, the reliability index needs to be normalized first to improve the accuracy of the calculation. This article uses a simple min-max normalization method [15] to scale the value of each element to the range [0], [1]. The normalized matrices are node CPU resource mapping experience matrix M'_{CPU}, node link resource mapping experience matrix M'_{LINK}, and node reliability experience matrix M'_{RELIAB}, respectively. The final underlying node reliability matrix M based on historical information is the sum of three matrices M'_{CPU}, M'_{LINK}, and M'_{RELIAB}, where the value of each element $m_{ij} \in M$ represents the reliability of the underlying node obtained based on the historical information. When $i = j$, $m_{ii} \in M$ represents the historical importance and reliability of the underlying node $n_i^s \in N_S$. When $i \neq j$, $m_{ij} \in M$ represents the historical correlation between the underlying node $n_i^s \in N_S$ and the underlying node $n_j^s \in N_S$. That is, after the virtual node $n_i^v \in N_V$ is mapped to the underlying node $n_i^s \in N_S$, the correlation between $n_i^s \in N_S$ and the underlying node $n_j^s \in N_S$, in which the virtual node $n_j^v \in N_V$ adjacent to the virtual node $n_i^v \in N_V$ is mapped on $n_j^s \in N_S$.

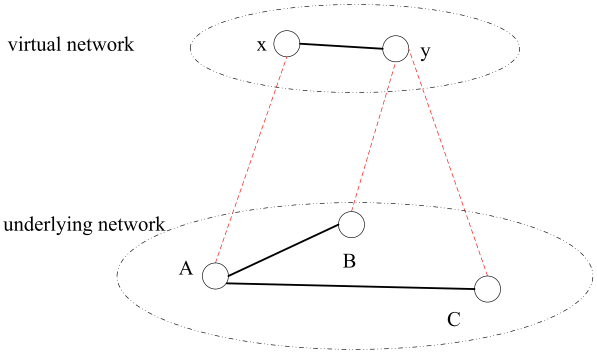
VI. BAYESIAN NETWORK BASED VIRTUAL NODE MAPPING MODEL

$N_V = \{n_i^v, n_2^v, \ldots, n_i^v, \ldots, n_m^v\}$ is used to represent a node set of virtual network containing m nodes, and $N_S = \{n_i^s, n_2^s, \ldots, n_i^s, \ldots, n_m^s\}$ is used to represent an underlying network node set mapped by m virtual network nodes. The probability of successful mapping of virtual node n_m^v can be calculated using formula (12). Among them, $P(n_m^s | n_1^s, \ldots, n_{m-1}^s)$ represents the probability of selecting n_m^s to map virtual node n_m^v on the premise that virtual nodes $n_1^v, \ldots, n_{m-1}^v$ have been successfully mapped.

$$P(n_m^v) = P(n_m^s | n_1^s, \ldots, n_{m-1}^s) \ldots P(n_2^s | n_1^s) P(n_1^s) \quad (12)$$

In the process of calculating the virtual network node mapping based on the historical node's reliability matrix of the underlying nodes, when the virtual network has only two nodes and the degree of the underlying network is small, the selection of the underlying nodes is relatively simple. However, as the scale of the virtual network increases and the scale of the underlying network increases, the reliability matrix becomes larger, and the problem of calculating virtual node mapping becomes more difficult. To solve this problem, this article proposes a reasoning model based on Bayesian networks [18]. Assuming that the mapping of the current virtual node is only related to its directly connected and mapped virtual node, formula (12) is simplified as shown in formula (13), where n_m^s represents the underlying node to which virtual node n_m^v will be mapped. $pa(n_m^v)$ represents the underlying node mapped by the parent node of virtual node n_m^v.

$$P(n_m^v) = P(n_m^s | pa(n_m^v)) \quad (13)$$



(a) Example of virtual network mapping

$$M = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix}$$

(b) Example of reliability matrix of underlying nodes based on historical information

FIGURE 2. Example of virtual network mapping.

To facilitate the understanding of the underlying node reliability matrix based on historical information, **Fig. 2**. is used as an example for illustration. The underlying network in **Fig. 2**. (a) includes three underlying nodes A, B, and C, and two underlying links, AB and AC. The reliability matrix M of the underlying node based on historical information is shown in **Fig. 2**. (b). After the virtual node x of the virtual network 1 has been mapped to the node A of the underlying network, find out which underlying node the virtual node y is mapped to. The following uses the underlying node reliability matrix M as an example for description. Assume that the underlying node B and the underlying node C are independent of each other. After the virtual node x of the virtual network 1 has been mapped to the node A of the underlying network, the virtual node y can be mapped to the underlying node B or C. According to the definition of $m_{ij} \in M$, it is known that the probability that the virtual node y can be mapped to the underlying node B is $P(B|A) = 4$. The probability that the virtual node y can be mapped to the underlying node C is $P(C|A) = 2$. Due to $P(B|A) > P(C|A)$, the virtual node y is mapped onto the underlying node B.

When the number of the underlying node $pa(n_m^v)$ mapped by the parent node of virtual node n_m^v is greater than or equal to 2, that is, $pa(n_m^v) = \{n_1^{pa,s}, n_2^{pa,s} \ldots, n_j^{pa,s}\}$, where $n_j^{pa,s}$ represents the j-th underlying node of the parent node of the virtual node n_m^v. Equation (13) becomes equation (14).

$$P(n_m^v) = P(n_m^s | n_1^{pa,s}, n_2^{pa,s} \ldots, n_j^{pa,s}) \quad (14)$$

According to conditional probability reasoning,

$$\begin{aligned} P(n_m^v) &= P(n_m^s | n_1^{pa,s}, n_2^{pa,s} \ldots, n_j^{pa,s}) \\ &= \frac{P(n_m^s, n_1^{pa,s}, n_2^{pa,s} \ldots, n_j^{pa,s})}{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_j^{pa,s})} \end{aligned} \quad (15)$$

When the j underlying nodes are independent of each other, according to Bayes' theorem,

$$\begin{aligned} P(n_m^v) &= \frac{P(n_m^s, n_j^{pa,s}) P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s} | n_m^s, n_j^{pa,s})}{P(n_j^{pa,s}) P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s})} \\ &= P(n_m^s | n_j^{pa,s}) \frac{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s} | n_m^s, n_j^{pa,s})}{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s})} \end{aligned}$$

Because the parent nodes are independent of each other, the above formula is:

$$P(n_m^v) = P(n_m^s | n_j^{pa,s}) \frac{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s} | n_m^s)}{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s})}$$

Because of

$$P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s} | n_m^s) = \frac{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s}, n_m^s)}{P(n_m^s)},$$

so, formula (15) can be simplified to formula (16).

$$\begin{aligned} P(n_m^v) = P(n_m^s | n_j^{pa,s}) &\times \frac{1}{P(n_m^s)} \\ &\times \frac{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s}, n_m^s)}{P(n_1^{pa,s}, n_2^{pa,s} \ldots, n_{j-1}^{pa,s})} \end{aligned} \tag{16}$$

Since the right part of formula (16) is similar to formula (15), therefore, formula (16) can be simplified to formula (17).

$$P(n_m^v) = \frac{1}{P(n_m^s)^{m-1}} \prod_{j=1}^{m} P(n_m^s | n_j^{pa,s}) \tag{17}$$

VII. HEURISTIC MAPPING ALGORITHMS

In order to make full use of the underlying node reliability matrix, based on network characteristics and node reliability matrix models, this article proposes two reliable virtual network mapping algorithms based on network characteristics and association relationships, namely the two-stage mapping algorithm for preferentially mapping virtual nodes (NFA-TS), and virtual network mapping algorithm based on hierarchical relationship (NFA-LR).

A. NFA-TS ALGORITHM

The algorithm NFA-TS is shown in **Table 1**. It mainly includes two stages: virtual node mapping and virtual link mapping. In the virtual node mapping phase, the importance $IMPORT(n_i^v)$ of the virtual node n_i^v is first calculated, and then the most important virtual node is allocated resources by the underlying node with the largest $m_{ii} \in M$. For other virtual nodes, use formula (17) to solve $P(n_i^v)$ and select underlying node, which satisfies $cpu(n_i^v)$ requires and has the largest $P(n_i^v)$, to allocate resources. In the virtual link mapping stage, the K-shortest path algorithm is used to allocate the underlying link resources for each virtual link that meet its constraints.

TABLE 1. NFA-TS algorithm.

```
Input: $G_S = (N_S, E_S)$ , $G_V = (N_V, E_V)$ , Node reliability matrix $M$ of the underlying network $G_S$
Output: Mapping list of $G_V$
1. For each virtual node $n_i^v \in N_V$ of the virtual network, use formula (9) to calculate the importance $IMPORT(n_i^v)$ of the virtual node;
2. Sort in descending order of $IMPORT(n_i^v)$ to obtain the virtual node set $N_V^{'}$ ;
3. For the largest virtual node of $IMPORT(n_i^v)$ in $N_V^{'}$ , select the underlying node that meets the needs of $cpu(n_i^v)$ and the largest $m_{ii} \in M$ from the set of underlying nodes to allocate resources; if the allocation is successful, delete the current virtual node $n_i^v$ from the set $N_V^{'}$ ; if the allocation fails, end;
4. Map each virtual node $n_i^v$ in $N_V^{'}$ in order;
    a) Use formula (17) to solve $P(n_i^v)$ , and assign the largest underlying node that meets $cpu(n_i^v)$ to the current virtual node;
    b) if the $cpu(n_i^v)$ of underlying node is not satisfied, the mapping fails and ends;
5. Allocate resources to each virtual link $e_j^v \in E_V$ in $E_V$ ;
    a) For each virtual link $e_j^v$ , use the k shortest path algorithm to find the underlying link that meets the link constraint $bw(e_j^v)$ , and allocate resources for $e_j^v$ ;
    b) If allocation fails, end.
```

B. NFA-LR ALGORITHM

In order to use the association relationship between virtual nodes, based on the algorithm NFA-TS, this article proposes the algorithm NFA-LR as shown in **Table 2**. The algorithm first calculates the importance $IMPORT(n_i^v)$ of the virtual node, sets the most important virtual node as the root node, and allocates resources using the underlying node that meets the $cpu(n_i^v)$ of node and has the largest $m_{ii} \in M$. A breadth-first search tree of virtual nodes is generated based on the root nodes of the virtual network. At each level of the breadth-first search tree, the resource allocation of virtual nodes and virtual links is achieved separately.

VIII. PERFORMANCE ANALYSIS

A. EXPERIMENTAL SETTINGS

In the existing research on virtual network resource allocation, the GT-ITM tool is used to generate the underlying network topology and virtual network mapping request [19]. In terms of generating the underlying network topology, there

TABLE 2. NFA-LR algorithm.

```
Input: $G_S=(N_S,E_S)$, $G_V=(N_V,E_V)$, Node reliability matrix $M$ of the underlying network $G_S$
Output: Mapping list of $G_V$
1. For each virtual node $n_i^v \in N_V$ of the virtual network, use formula (9) to calculate the importance $IMPORT(n_i^v)$ of the virtual node;
2. Select the virtual node $n_i^v$ with the largest $IMPORT(n_i^v)$ as the root node;
3. From the set of underlying nodes, select the underlying node that meets the needs of $cpu(n_i^v)$ and the largest $m_{ii} \in M$ to allocate resources to the root node. If the allocation fails, the process ends.
4. Perform a breadth-first search on the virtual network based on the root node to generate a breadth-first search tree $Tree(n_i^v)$;
5. Traverse virtual nodes of each layer of $Tree(n_i^v)$:
a) Sort the nodes of this layer in descending order of $IMPORT(n_i^v)$;
    i. Use formula (17) to solve $P(n_m^v)$, and assign resource of underlying node to the current virtual node that meets $cpu(n_i^v)$'s needs and has the largest $P(n_m^v)$;
    ii. If the allocation fails, end;
b) The virtual link set corresponding to the current virtual node is arranged in descending order of $bw(e_j^v)$ and mapped one by one;
    i. Use the k shortest path algorithm to allocate the lowest link resources for the current virtual link with the least number of hops;
    ii. If the allocation fails, determine whether the number of backtracking nodes exceeds the backtracking threshold $\eta$. If it exceeds, the mapping fails; otherwise, it traces back to the previous virtual node and remaps the underlying nodes for it.
```

are 100 underlying nodes, and any two underlying nodes are connected with a probability of 0.5. This underlying network topology is equivalent to the network resources provided by a medium-sized network provider. In terms of virtual network topology generation, virtual nodes obey the uniform distribution of [2], [8], and any two virtual nodes are connected with a probability of 0.5. In terms of resource assignment, set the value range of the CPU resource value and link bandwidth resource value of the underlying node and the underlying link to a uniform distribution of [30], [60]. Because the value ranges of node resources and link resources are the same, set the value of κ and λ parameters to 1. Similar to [6], this article sets the traceback threshold η to 3n, where n is the number of virtual network nodes. The CPU resource value range of the virtual node is set to a uniform distribution of [1], [5]. The value range of the virtual link bandwidth resource value is set to a uniform distribution of [1], [10]. A total of 6000-time units were run in the experiment. As for the virtual network mapping request, the arrival time of the virtual network mapping request is set to obey a Poisson distribution with an average interval of 1.5-time units. It contains about 4000 virtual network mapping requests. The average life cycle of each virtual network request is 20-time units. The experimental hardware platform is configured as a cloud host with 8-core CPU, 8G memory, and 200G hard disk, and the operating system is CentOS 6.6.

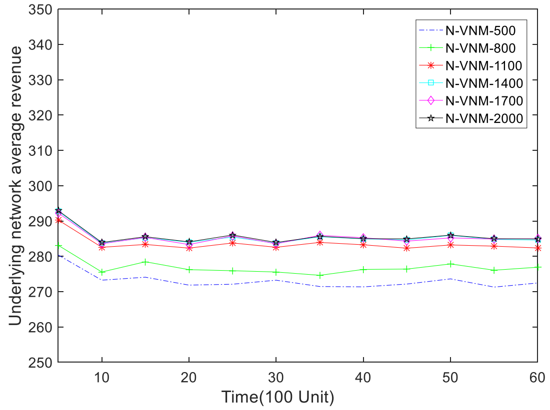


FIGURE 3. Underlying network average revenue.

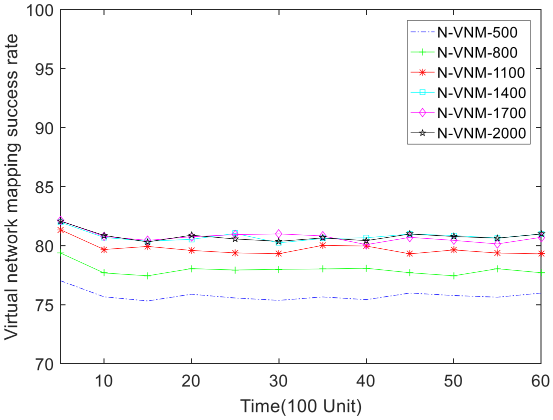


FIGURE 4. Virtual network mapping success rate.

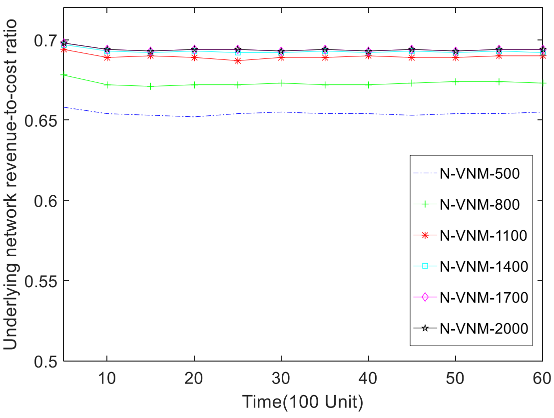


FIGURE 5. Underlying network revenue-to-cost ratio.

B. NODE RELIABILITY MATRIX MODELING ANALYSIS

In order to analyze the impact of the constructed node reliability matrix on the performance of the algorithm, this section analyzes the network revenue, mapping success rate, revenue-cost ratio of underlying layer of the algorithm NFA-TS when the node reliability matrix is constructed under different numbers of virtual network mappings (N-VNM). The experimental results are shown in **Fig. 3-5**., comparing the impact of the node reliability matrix constructed on the N-VNM values of 500, 800, 1100, 1400, 1700, and 2000 on the performance of the algorithm NFA-TS. From the experimental results, it can be seen that with the increase of the value of N-VNM, the performance of the algorithm NFA-TS's underlying network revenue, mapping success rate, and revenue-to-cost ratio gradually increases. When the value of N-VNM increases to 1400, the performance of the algorithm NFA-TS's underlying network revenue, mapping success rate, and revenue-to-cost ratio tends to stabilize. The experimental results show that when the number of N-VNM is small, the node reliability matrix does not reflect the network characteristics and association relationships of the underlying network, so the performance of the algorithm NFA-TS is low. When the number of N-VNM is large, the node reliability matrix can reflect the network characteristics and association relationships of the underlying network, so the performance of the algorithm NFA-TS is gradually improved to become stable.

C. COMPARATIVE ANALYSIS WITH RELATED ALGORITHMS

The performance of the proposed algorithms NFA-TS and NFA-LR is analyzed using the node reliability matrix constructed when the N-VNM value is 1400.

It can be known from the analysis of existing researches that the current virtual network mapping algorithm implementation can be divided into two types: heuristic algorithm and mathematical planning algorithm. Among them, the time cost of mathematical planning algorithm increases exponentially with the increase of the number of nodes in the underlying network. This article uses a heuristic algorithm, so we need to choose a heuristic algorithm that solves similar problems. Because the algorithm SVNE [10] and the algorithm VNE-SSM [14] are similar to the problem solved by the algorithm in this article, and they are more classic algorithms. In the following, the algorithm shown in **Table 3** is analyzed from five aspects, such as the average cost and revenue of the underlying network, the success rate of virtual network mapping, and the average utilization of the underlying network links and nodes. The experimental results are shown in **Fig. 6-10**.

TABLE 3. Algorithm name and description.

Name	Description
NFA-TS	A two-stage mapping algorithm for priority mapping virtual nodes proposed in this paper;
NFA-LR	Virtual network mapping algorithm based on hierarchical relationship proposed in this paper;
SVNE	Add the parameter values that affect the reliability of the underlying network to obtain the underlying network resources with higher reliability, and propose a reliability-aware virtual network mapping algorithm [10];
VNE-SSM	Use the greedy algorithm to maximize the utilization of the underlying network resources [14];

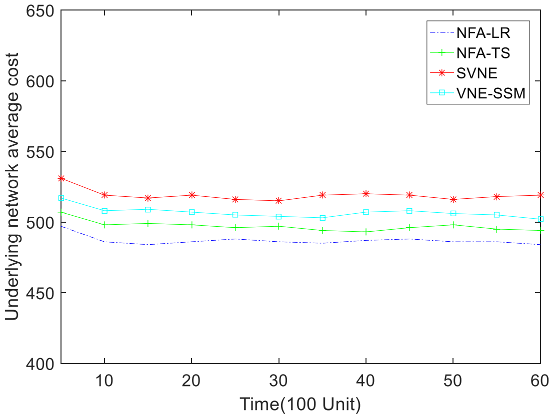


FIGURE 6. Analysis of the average cost of the underlying network.

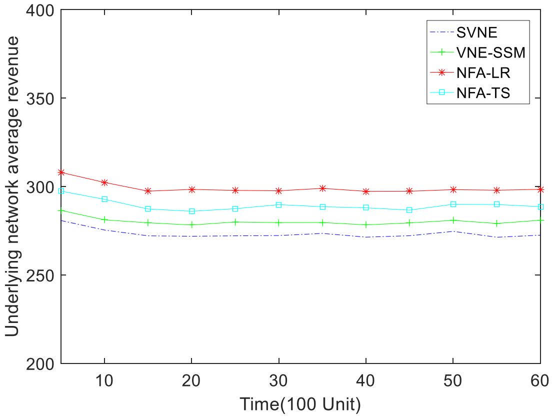


FIGURE 7. Analysis of the average revenue of the underlying network.

The analysis of the average cost of the underlying network is shown in **Fig. 6**. The average cost of the underlying network of the algorithm SVNE is the largest. The average network cost of the algorithm NFA-LR and NFA-TS proposed in this article is lower, which indicates that the algorithm in this article allocates underlying network link resources with fewer hops to the virtual link. The average revenue of the underlying network is shown in **Fig. 7**. The average revenue of the underlying network of the algorithm NFA-LR is the largest. It shows that the algorithm in this article allocates the underlying network resources for more virtual links and obtains more virtual network benefits. The virtual network mapping success rate is shown in **Fig. 8**. The algorithm NFA-LR's virtual network mapping success rate is the largest, indicating that the algorithm in this article receives more virtual network resource allocation requests and more virtual network requests are successfully allocated resources. The average utilization rate of the underlying network link is shown in **Fig. 9**. The average utilization rate of the underlying network link of the algorithm NFA-LR is the largest, indicating that the algorithm in this article uses more resources of the underlying network link. The average utilization of the underlying network nodes is shown in **Fig. 10**. The average

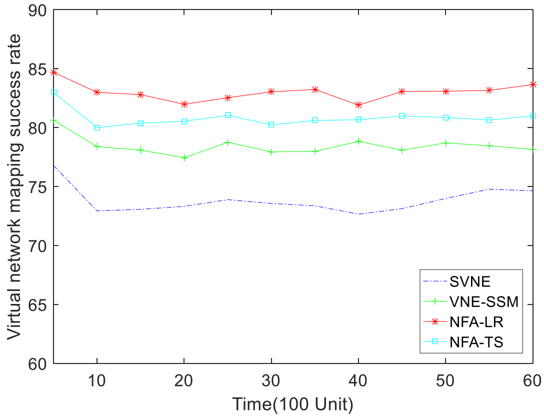


FIGURE 8. Virtual network mapping success rate analysis.

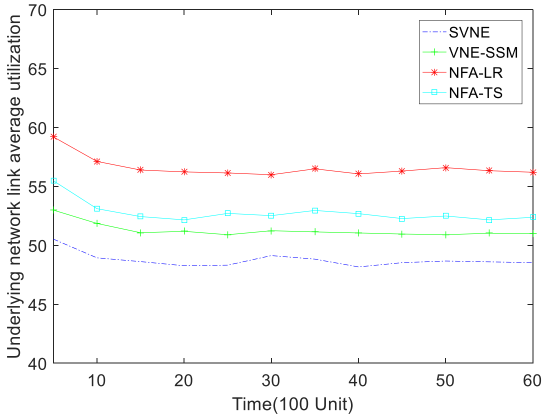


FIGURE 9. Underlying network link average utilization analysis.

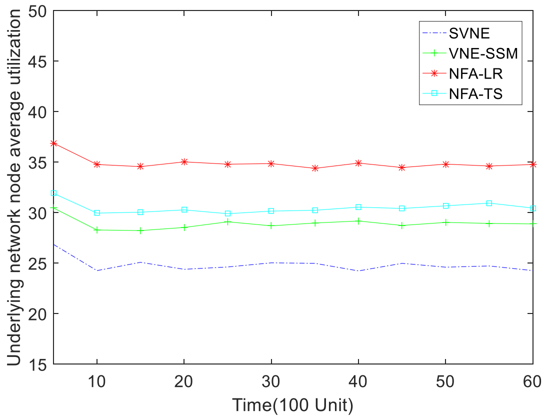


FIGURE 10. Underlying network node average utilization analysis.

utilization of the underlying network nodes of the algorithm NFA-LR is the largest, indicating that the algorithm in this article uses more resources of the underlying network nodes.

From the experimental results, compared with the existing research, the algorithm in this article has achieved better performance indicators in five aspects, such as the average cost and benefits of the underlying network, the success rate of virtual network mapping, and the average utilization of the underlying network links and nodes. Therefore, the algorithm proposed in this article can make full use of the network characteristics and association relationships of the underlying network by analyzing a large amount of historical data, and select the appropriate underlying network resources to achieve the shortest link mapping effect, effectively avoiding the existing algorithms to preferentially use the underlying network resources with the largest resources, and reduce the fragmentation of the underlying network resources. In addition, the hierarchical relationship of the algorithm NFA-LR uses the association relationship between virtual nodes for virtual network mapping. Compared with the algorithm NFA-TS, it obtains a more optimized underlying network resource, which further improves the utilization of the underlying network resource and virtual network mapping success rate.

IX. CONCLUSION

In recent years, network virtualization has become a key technology to solve the problem of network rigidity. Virtual network resource allocation has become a key research hotspot in network virtualization. In the existing research on reliable virtual network mapping, the main problems include: (1) In the virtual node mapping phase, the key parameters of the underlying nodes are simply added, and the largest underlying node is used to allocate resources, which is not conducive to reflecting the characteristics of the network. It is also easy to cause fragmentation of the underlying resources; (2) In the virtual network mapping, only the current virtual network request and the underlying network resource status are considered, and the network topology and historical mapping data are not fully utilized. In order to solve these problems, this article combs the network characteristics related to reliable virtual network mapping, establishes the underlying node reliability matrix based on historical information and the Bayesian network-based inference model, and proposes a reliable virtual network mapping algorithm based on network characteristics and association relationships. The algorithm is compared with related classical algorithms through simulation experiments. It is verified that the algorithm in this article is superior to the existing virtual network mapping algorithms in terms of virtual network mapping success rate and the utilization of the underlying network resources.

In order to ensure that the virtual network can survive and improve the survivability of the virtual network through migration and reconfiguration technology when the underlying network fails, data mining and machine learning technology will be used in the next work based on the research results of this article to further mine the migration and reconfiguration association relationship between the underlying network resources when the underlying network fails, so as to achieve a highly survivable virtual network mapping algorithm.

REFERENCES

[1] N. M. M. K. Chowdhury and R. Boutaba, "Network virtualization: State of the art and research challenges," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 20–26, Jul. 2009.

[2] A. Fischer, J. F. Botero, B. M. Till, H. de Meer, and X. Hesselbach, "Virtual network embedding: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1888–1906, 4th Quart., 2013.

[3] M. Melo, S. Sargento, U. Killat, A. Timm-Giel, and J. Carapinha, ''Optimal virtual network embedding: Node-link formulation,'' *IEEE Trans. Netw. Service Manage.*, vol. 10, no. 4, pp. 356–368, Dec. 2013.
[4] R. Mijumbi, J. Serrat, J.-L. Gorricho, and R. Boutaba, ''A path generation approach to embedding of virtual networks,'' *IEEE Trans. Netw. Service Manage.*, vol. 12, no. 3, pp. 334–348, Sep. 2015.
[5] P. Zhang, H. Yao, M. Li, and Y. Liu, ''Virtual network embedding based on modified genetic algorithm,'' *Peer Peer Netw. Appl.*, vol. 12, no. 2, pp. 481–492, Mar. 2019.
[6] C. K. Dehury and P. K. Sahoo, ''DYVINE: Fitness-based dynamic virtual network embedding in cloud computing,'' *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1029–1045, May 2019.
[7] M. Dolati, S. B. Hassanpour, M. Ghaderi, and A. Khonsari, ''DeepViNE: Virtual network embedding with deep reinforcement learning,'' in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2019, pp. 879–885.
[8] A. Jahani, L. M. Khanli, M. T. Hagh, and M. A. Badamchizadeh, ''Green virtual network embedding with supervised self-organizing map,'' *Neurocomputing*, vol. 351, pp. 60–76, Jul. 2019.
[9] A. Jahani, L. M. Khanli, M. T. Hagh, and M. A. Badamchizadeh, ''EE-CTA: Energy efficient, concurrent and topology-aware virtual network embedding as a multi-objective optimization problem,'' *Comput. Standards Interfaces*, vol. 66, pp. 1–17, Oct. 2019.
[10] M. R. Rahman and R. Boutaba, ''SVNE: Survivable virtual network embedding algorithms for network virtualization,'' *IEEE Trans. Netw. Service Manage.*, vol. 10, no. 2, pp. 105–118, Jun. 2013.
[11] M. M. A. Khan, N. Shahriar, R. Ahmed, and R. Boutaba, ''Multi-path link embedding for survivability in virtual networks,'' *IEEE Trans. Netw. Service Manage.*, vol. 13, no. 2, pp. 253–266, Jun. 2016.
[12] S. R. Chowdhury, R. Ahmed, M. M. A. AKhan, N. Shahriar, R. Boutaba, J. Mitra, and F. Zeng, ''Dedicated protection for survivable virtual network embedding,'' *IEEE Trans. Netw. Service Manage.*, vol. 13, no. 4, pp. 913–926, Dec. 2016.
[13] H. Jiang, Y. Wang, L. Gong, and Z. Zhu, ''Availability-aware survivable virtual network embedding in optical datacenter networks,'' *J. Opt. Commun. Netw.*, vol. 7, no. 12, pp. 1160–1171, 2015.
[14] V. Lira, E. Tavares, M. Oliveira, Jr., E. Sousa, and B. Nogueira, ''Virtual network mapping considering energy consumption and availability,'' *Computing*, vol 101, no. 8, pp. 937–967, 2019.
[15] X. Zheng, J. Tian, X. Xiao, X. Cui, and X. Yu, ''A heuristic survivable virtual network mapping algorithm,'' *Soft Comput.*, vol. 23, no. 5, pp. 1453–1463, Mar. 2019.
[16] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, nos. 1–2, pp. 5–43, Jan. 2003.
[17] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2006.
[18] R. E. Neapolitan, *Learning Bayesian Networks*. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[19] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee, ''How to model an Internetwork,'' in *Proc. IEEE Infocom*, Mar. 1996, pp. 594–602.

SHUNLI ZHANG received the M.S. degree in computer science from the Taiyuan University of Technology, China, and the Ph.D. degree in computer science from the Beijing University of Posts and Telecommunications in 2012. He is currently a Postdoctoral Researcher and a Senior Engineer with Jinzhong University, Shanxi, China. He has authored journal and conference papers on resource management and fault diagnosis of network. His research focuses on management for virtual networks and next generation networks.

• • •