研究与开发

基于聚类的安全分级虚拟网络映射方法

陆勰，徐雷，张曼君
（中国联合网络通信集团有限公司研究院，北京 100048）

摘　要：基于 5G 网络环境，提出了一种基于 *K*-means 聚类的安全分级虚拟网络映射方法，特别涉及一种对节点安全分级的思想，结合 *K*-menas 聚类算法，解决当下虚拟网络映射存在的低安全性、低映射效率等关键问题，与传统方法相比，所提方法提高了虚拟节点映射的效率，增强了网络韧性。

关键词：聚类；安全分级；虚拟节点；映射；*K*-means
中图分类号：TP393
文献标识码：A
doi: 10.11959/j.issn.1000−0801.2021224

Safe hierarchical virtual network mapping method based on clustering

LU Xie, XU Lei, ZHANG Manjun
Research Institute of China United Network Communications Group Co., Ltd., Beijing 100048, China

Abstract: Based on the 5G network environment, a security hierarchical virtual network mapping method based on *K*-means clustering was proposed. In particular, the idea of node security classification and the combination of *K*-menas clustering algorithm was involved to solve the key problems of low security and low mapping efficiency existing in current virtual network mapping. Compared with traditional methods, this research method improves the efficiency of virtual node mapping and enhances the network resilience.

Key words: clustering, security classification, virtual node, mapping, *K*-means

1　引言

1.1　虚拟网络映射的概念

随着 5G 网络的快速发展，三大应用场景增强型移动宽带（enhanced mobile broad，eMBB）、低时延高可靠通信[1]（ultra-reliable and low-latency communication，uRLLC）以及大连接物联网（massive machine-type communication，mMTC）与垂直行业的深度融合，对网络需求不断升级，为了更好地服务行业与用户，5G 网络软、硬件分离，SDN 与 NFV 技术顺势而出，5G 网络虚拟化提高了网络的灵活性、多样性，释放了更多网络资源的能力，但是由于底层的物理基础设施是共享的，当其中一个物理服务器发生故障或受到攻击时会导致其上所承载的虚拟机失效，影响网络与业务的连续性。针对虚拟化带来的安全风险，虚拟网络映射作为解决虚拟化安全的重要关键技术，成为当下领域研究的热点。虚拟网络由虚拟

收稿日期：2021−04−16；修回日期：2021−05−25

节点与虚拟链路的映射两部分组成，为虚拟网络分配底层物理资源称为虚拟网络映射[2]，虚拟网络映射主要解决虚拟节点到物理节点的映射以及虚拟链路到对应的物理链路的路径问题。

1.2　虚拟网络映射研究现状

面对虚拟化网络映射问题，现有的解决措施是可生存性的虚拟网映射[3]（survivable virtual network embedding，SVNE），也就是对异常的虚拟机节点备份和重映射。备份主要针对对网络有高要求的环境，网络节点需要有备份策略，该种方式资源占用较大、成本较高，如高需求的网络切片[4]。而重映射属于针对动态网络环境下采取的映射方法相对的备份方式，资源灵活度高、成本较低，因此重映射的适用场景较广。就目前来看重映射采取的具体映射方法，主要围绕系统的特性、节点的重要程度、节点映射的顺序以及节点之间的信任度等方向。如季新生[4]等提出一种基于操作系统异构的虚拟网映射方法，该种方法主要侧重虚拟网络节点与物理服务器的系统异构性，如果虚拟节点与物理服务器的操作系统一样，恶意攻击者会通过探测原物理服务器的系统类型，挖掘系统漏洞从而发动攻击，促使虚拟机请求映射，如果映射到的物理服务器也是同系统，则默认具有同样的漏洞，这样攻击者就能轻而易举地再次攻击对映射后的物理服务器。这种方法较局限，只考虑到系统的异构带来的风险，未能从全局出发，考虑节点其他属性的安全性，无法从根本上达到映射的目的。刘新波等[5]提出一种基于多属性节点重要性排序的映射方法，该方法主要侧重考虑虚拟节点的资源能力，如主要考虑自身 CPU 资源和邻近的链路资源，按照重要性排序并以此为标准来寻找符合的物理机，这样的后果是节点排序未考虑其拓扑结构，容易造成虚拟节点映射到较远的物理机上，导致映射效率低下、资源浪费。龚水清[6]等提出一种基于信任度的安全虚拟网络映射，该方法是对虚拟节点与物理节点之间的信任评估，确保映射发生在二者相互信任的情况下，其次，按照节点属性的重要程度排序，保障较为重要的虚拟节点能够映射到较为重要的物理机上，同时采用 K 最短路径算法完成链路映射，该方法一定程度上把安全作为映射的基础，根据 CPU 约束及节点位置约束筛选物理节点集合，根据信任度选择最佳物理节点，但该方法没有减少链路映射时 K 最小路径算法的迭代次数，且节点映射的算法先在全网筛选符合条件的物理节点集合，算法的复杂度较高。

1.3　现有技术缺陷

综上，以上研究主要考虑映射发生时采取的解决方案，未从全局出发，采取的解决方法局限在某一方面，很多没有重点考虑映射最重要的基础与目标，故以上研究存在着两方面的缺陷，首先，没有把节点的安全性作为一个映射属性加以考虑，因为当出现把一个安全级别较高的虚拟节点映射到较低安全级别的物理节点上时，受到攻击的风险加大，尤其是处于网络核心的节点一旦遭受攻击，后果不堪设想；其次，映射效率低下，现有算法的映射从全网视角，虚拟节点映射需要遍历整个网络中的物理节点，挑选符合映射条件的节点，再映射链路，链路映射通常利用 K 最短路径算法找到最短路径，当业务发生时，这种方法会耗费大量的时间寻找符合要求的物理节点，且时延作为检验网络性能的关键指标，如果映射时间过长，导致时延的增加，就会造成业务“失联”、网络拥塞等重大问题。

1.4　基于聚类的安全分级虚拟网络映射方法的思想及意义

为了克服上述现有技术的不足，本文提出了一种基于聚类的安全虚拟网络映射方法。总体思想主要包括 3 个方面，首先，对虚拟和物理节点量化、安全分级，保障虚拟节点与物理节点的映射是同安全等级，然后利用 K-means[7]聚类完成相似程度最大的节点分类，摸清网络节点安全现状；

其次，当异常事件发生时，利用事前掌握的网络节点情况，与欧氏距离[8]相结合寻找异常节点最佳映射节点，完成事中异常虚拟节点安全映射，最后利用 *K* 最短路径算法[9]找出最短路径完成链路映射，可依据此业务链完成事后对异常事关联分析、追踪溯源。该方法具有两方面的重要意义，一方面通过节点安全分级、聚类等方式缩小了映射范围，提高了映射的效率与成功率；另一方面通过欧氏距离计算同一类簇中虚拟节点与物理节点的距离，减少了链路算法迭代的次数，降低了算法的复杂度[10]。

2 基于聚类的安全分级虚拟网络映射方案

2.1 系统架构

按照本文提出的对虚拟节点的安全分级与映射前聚类、同类簇中利用欧氏距离寻找最佳映射的物理节点，最后利用最短路径算法完成链路映射的思想，所提出的原型系统框架如图 1 所示。

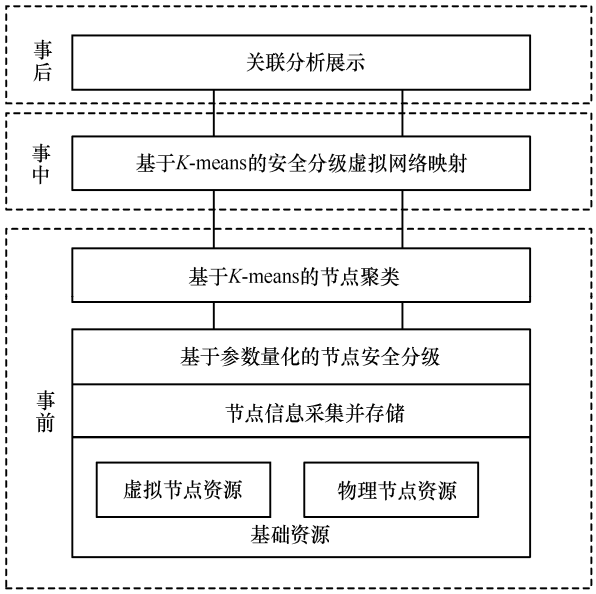


图 1 基于 *K*-means 聚类的安全分级虚拟网络映射方法的原型系统框架

图 1 中事前、事中和事后 3 个阶段纵向联动，完成对异常节点的映射。具体来说，事前做到对各节点“心中有数”，通过建立量化模型，把各节点的属性参数值进行量化、安全分级，然后进行 *K*-means 聚类，找出各节点间的相似性，形成 *K* 个类簇，并对各个节点进行类簇标识，形成各节点的键值对；事中快速响应，当虚拟节点发生异常需要映射时，优先从与该节点处于同一类簇中的物理节点集合中选择最佳节点（查看标识很快筛选出相应的类簇），因为同一簇中的节点具有更大的相似性，这样的映射能更好地保障网络运行的可靠性与稳定性，且是对资源的合理有效利用，这样的映射类似“假备份”，在同一类簇中利用欧氏距离算法选择符合要求的物理节点，再利用 *K* 最短路径算法找出最短路径完成链路映射；事后定位溯源，由于前期做了聚类标识，当出现故障时，能够快速定位故障点，各节点关联分析，缩短故障排查时间、提高运维效率。

2.2 虚拟节点映射流程

按照系统的总体框架，虚拟节点映射系统流程如图 2 所示，主要有以下 4 个步骤。

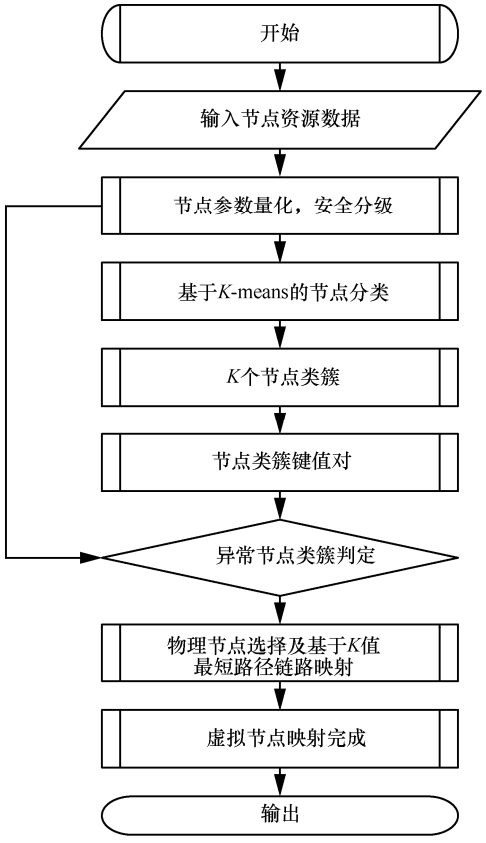


图 2 基于聚类的安全分级虚拟节点映射系统流程

（1）采集节点资源数据

通过采集软件或采集器等采集或对接数据库读取网络中的虚拟节点和物理节点的基本信息，包括 CPU、存储、带宽、位置、功能等关键信息，具体读取哪些属性参数可依据自身网络环境的侧重点有所调整，信息采集完毕后存储。

（2）节点参数量化、安全分级

传统的网络没有把内生安全作为考察网络鲁棒性的重要指标，结合本文研究的重点，无论是虚拟节点还是物理节点，需要对它进行安全分级，事先把安全“内生”于网络环境，减少使用“亡羊补牢”或“烟囱式”的安全手段，虽然安全分级自身没有绝对的标准，但最终目标都是为了保障网络的韧性和业务的可靠性。

首先建立一个安全可行的评估模型，针对步骤（1）中的节点信息，分别对各个属性进行量化评估，主要目的是计算每一个节点的安全等级，采取的分级标准主要依据节点 CPU 大小、位置、所链接链路的带宽之和、物理资源需求等，节点的属性参数可表示虚拟节点 $n_{vi}=\{c_{vi},l_{vi},b_{vi},r_{vi}\}$，物理节点 $n_j=\{c_j,l_j,b_j,r_j\}$，节点集合 $n=\{n_{vi},n_j\}$，其中，c 表示为 CPU 大小，l 表示位置，b 表示节点所链接的链路带宽总和，r 表示其他一些关键资源，i 的最大值是网络中虚拟节点的数量，j 的最大值是网络中物理节点数量（以下涉及的地方均是相同定义范围）。安全需求的评估本身没有一个绝对的界限与标准，可根据各自所处的环境结合实际建立恰当的评估模型，本评估模型系统中对虚拟节点的安全需求和物理节点的安全作出相应的评估，并给出合理的安全评分，且评分满足均匀分布，节点安全值等于各个属性对应参数值量化评估后的总和，各个属性评分范围限定在（0,1），如 CPU 的评分范围是（0,0.3），针对节点 CPU 的大小，划定评估线，如 CPU 主频是[2,3] GHz 评分 0.2，低于[2,3] GHz 评分 0.1，高于[2,3] GHz 评分 0.3 等；节点位置评分范围是（0,0.4），根据节点位置给予评分，如根据网络域的划分，处于安全域的评分高于非安全域，如果没有划分网络域的网络，可按照网络节点处的位置评分，如处于控制面的节点安全性高于数据面的节点安全性；节点链路带宽值评分范围是（0,0.6）等，通过划定各个量化参数值完成对各个参数的安全量化评估，最后节点的安全值 s 为各个属性对应参数值量化评估总和，如虚拟节点安全值 $s_{v1}=s_{vc1}+s_{vl1}+s_{vb1}+s_{vr1}$，物理节点安全值 $s_1=s_{c1}+s_{l1}+s_{b1}+s_{r1}$，节点集合的安全值集合记作 $S\{s_{v1},s_{v2},s_{v3},\cdots,s_{vi},s_1,s_2,s_3,\cdots,s_j\}$，根据 S 集合中各个安全值的分布特点，确定安全划线参考参数 d，如 $d=1$，s_{vi} 或 $s_j>d$ 为高安全等级，s_{vi} 或 $s_j=d$ 为中安全等级，s_{vi} 或 $s_j<d$，$i=[1,n]$为低安全等级，把 n 中的所有节点按照划分原则划定为高、中、低 3 个等级，分别使用 s_h、s_m 及 s_l 表示 3 个等级，量化数值采用 3、2、1，即 s_h 量化值是 3，为高安全等级，s_m 量化值是 2，为重安全等级，s_l 量化值是 1，为低安全等级，且把相应的等级标识添加到各个节点的参数集合中，即 $n_{vi}=\{s_w,c_{vi},l_{vi},b_{vi},r_{vi}\}$，物理节点 n_j 表示为 $n_j=\{s_w,c_j,l_j,b_j,r_j\}$，更新节点集合为 $n=\{n_{vi},n_j\}$，其中，$w=$[h,m,l]，也就是说，每一个虚拟节点和物理节点中均包含了安全等级和各量化的属性值。

（3）基于 *K*-means 算法的节点聚类

在步骤（2）完成后，对所有的节点，根据 *K*-means 聚类思想按照距离的远近划分，输入 k 值和 k 个初始聚类中心，对于节点集合 $n=\{n_{vi},n_j\}$，其中每一个节点都是一个多维的属性集合，按照 *K*-means 算法原理，采用欧氏距离计算两个节点间距离，经过多次迭代计算把 n 个节点按照距离聚类中心的远近划分到 k 个类簇中，标识为 $k=\{k_1,k_2,\cdots,k_u\}$，$u=[1,k]$，并把标识添加到各节点参数集合中，把每一个节点写成键值对后存储，其键值表达式为 $N_{vi}=<k_u,n_{vi}>$和 $N_j=<k_u,n_j>$，其中，$u=[1,k]$。

（4）物理节点选择及 K 值最短路径链路映射

当某个虚拟节点发生异常请求映射，异常情况包括但不限于故障、受到攻击，根据步骤（3）中最后得到的键值对中的 k 值查出异常节点所处的类簇，然后根据如下原则筛选出符合要求的物理节点，完成虚拟节点 $N_{vi}=<k_u,n_{vi}>$ 到 $N_j=<k_u,n_j>$ 的映射。

- 物理节点与虚拟节点同安全等级。
- 物理节点所承载的虚拟节点的带宽之和不超过该物理节点的总带宽。
- 物理节点相关链路不通过虚拟节点所对应的物理服务器，目的是避免同时失效。
- 同一虚拟网中的不同虚拟机不能映射到同一物理服务器上。

经过筛选形成候选物理节点集合，即 $G=\{n_1,n_2,\cdots,n_t\}$，其中，t 的值小于或等于该类簇（异常节点所在类簇）中物理节点数量，然后根据欧氏距离计算异常虚拟节点与集合 G 中各物理节点的距离，也就是寻找进一步的最大相似性，选择欧氏距离最小值对应的物理节点作为最佳映射的节点，完成异常虚拟节点到最佳物理节点的映射。传统模式下，异常虚拟节点的映射需要从全网络寻找合适的物理节点，这种模式无疑会增加网络响应时间、提高业务时延，而本文提出的方法通过筛选得到了缩小范围的物理节点集合 G，可极大缩短网络响应时间，且由于 G 中节点之间的相似性提高了映射的成功率与网络运行的稳定性，最大限度地保障了资源的合理利用，这种“假备份”的方式增加了网络的韧性。虚拟网络的映射包括节点和链路的映射。最后，根据链路映射算法原理，完成链路映射，根据 K 值最短路径算法，按照路径跳数大小升序排列，选出符合要求的最短路径，完成整个虚拟节点到物理节点的可靠映射。

2.3 性能分析

本节主要针对在异常虚拟节点映射时采取的方法进行复杂度分析，选取一个有 n 个虚拟节点，m 个物理节点的网络，在传统模式下，n 个虚拟节点映射的算法复杂度是 $O(nm^2)$，但如果使用本文提出的方法，在 m 个物理节点中，安全分级确定了同等级的虚拟节点与物理节点，聚类确保了在同一类簇中虚拟节点与物理节点的最大相似性，欧氏距离计算进一步提高了异常节点与物理节点的相似性，当使用 K 最短路径算法寻找映射链路时，缩小了链路选择的范围，减少了路径迭代算法的次数，因此，经过筛选得出的 t 个符合映射要求的物理节点一定小于 m，即 $t<m$，改进后的算法复杂度是 $O_1(nt^2)$，很明显 $O_1(nt^2)<O(nm^2)$。得出经过本文提出的基于安全分级聚类的虚拟节点映射方法，相比于传统方法，映射计算复杂度低、映射效率高。

通过本方法事前、事中的协作运行，为事后的关联分析提供了可靠的支撑，由于经过虚拟网络节点的映射，在各个节点上都添加了聚类标识，形成键值对，当网络中出现异常时，可根据聚类进一步进行关联分析，如当多个虚拟节点同时发起映射请求时，可通过查看各个虚拟节点所在类簇，寻找之间的关联关系，定位异常点，迅速排除故障。

3 结束语

本研究方法相比较传统的映射策略，首先，把安全作为一个重要的映射首要条件，增强了网络的鲁棒性；其次，从映射的效率来看，传统的虚拟节点映射时，对于一个虚拟节点数量为 n，物理节点为 m 的网络，映射时的计算复杂度是 $O(nm^2)$，而经过本研究中先聚类后映射的思想，对于虚拟节点数量 n 的集合，优先选择同类簇中的物理节点映射，降低了算法的复杂度，提升了映射的效率，更好地保障了业务的连续性。该研究方法的特点在于在业务发生前也就是事前对虚拟节点和物理节点的安全性和相关性进行了分

析，做到“心中有数”，当产生业务时，能够按照事前的分析，快速做出响应，提高了业务的连续性，同时增强了网络韧性，事后还可根据聚类进行关联分析，为运维提供可靠支撑。

5G 网络的高速发展，千行百业对网络需求差异化、精细化要求越来越高，伴随的安全差异化需求也不断提高，5G 网络切片技术、边缘云技术等新技术的广泛应用将会加大 5G 网络各方面的需求，而虚拟化作为 5G 网络的典型特征，虚拟化网络的安全将是保障各种新技术、新应用、新业务等的基础支撑，只有不断筑牢虚拟网络安全基础防线，提高网络映射的成功率与速率，以此增强网络韧性，才能更大地释放 5G 网络的“能量”，更好地服务用户、服务行业。

参考文献：

[1] 梁辉, 韩潇, 李福昌. 5G URLLC 无线网络部署方案分析[J]. 邮电设计技术, 2020(3): 15-18.
LIANG H, HAN X, LI F C. Analysis on 5G uRLLC wireless network deployment scheme[J]. Designing Techniques of Posts and Telecommunications, 2020(3): 15-18.

[2] 程祥, 张忠宝, 苏森, 等. 虚拟网络映射问题研究综述[J]. 通信学报, 2011, 32(10): 143-151.
CHENG X, ZHANG Z B, SU S, et al. Survey of virtual network embedding problem[J]. Journal on Communications, 2011, 32(10): 143-151.

[3] 肖蔼玲, 王颖, 孟洛明, 等. 面向多节点故障的生存性虚拟网络映射方法[J]. 通信学报, 2015, 36(4): 85-92.
XIAO A L, WANG Y, MENG L M, et al. Virtual network embedding approach to survive multiple node failures[J]. Journal on Communications, 2015, 36(4): 85-92.

[4] 季新生, 赵硕, 艾健健, 等. 异构备份式的虚拟网映射方法研究[J]. 电子与信息学报, 2018, 40(5): 1087-1093.
JI X S, ZHAO S, AI J J, et al. Research on heterogeneous-backup virtual network embedding[J]. Journal of Electronics & Information Technology, 2018, 40(5): 1087-1093.

[5] 刘新波, 王布宏, 刘帅琦, 等. 安全虚拟网络映射的启发式算法[J]. 系统工程与电子技术, 2018, 40(3): 676-681.
LIU X B, WANG B H, LIU S Q, et al. Heuristic algorithm for secure virtual network embedding[J]. Systems Engineering and Electronics, 2018, 40(3): 676-681.

[6] 龚水清, 陈靖, 黄聪会, 等. 信任感知的安全虚拟网络映射算法[J]. 通信学报, 2015, 36(11): 180-189.
GONG S Q, CHEN J, HUANG C H, et al. Trust-aware secure virtual network embedding algorithm[J]. Journal on Communications, 2015, 36(11): 180-189.

[7] 张冬梅, 李敏, 徐大川, 等. *K*-均值问题的理论与算法综述[J]. 中国科学: 数学, 2020, 50(9): 1387-1404.
ZHANG D M, LI M, XU D C, et al. A survey on theory and algorithms for *K*-means problems[J]. Scientia Sinica (Mathematica), 2020, 50(9): 1387-1404.

[8] 王子龙, 李进, 宋亚飞. 基于距离和权重改进的 *K*-means 算法[J]. 计算机工程与应用, 2020, 56(23): 87-94.
WANG Z L, LI J, SONG Y F. Improved *K*-means algorithm based on distance and weight[J]. Computer Engineering and Applications, 2020, 56(23): 87-94.

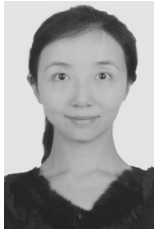
[9] 徐涛, 丁晓璐, 李建伏. *K* 最短路径算法综述[J]. 计算机工程与设计, 2013, 34(11): 3900-3906, 3911.
XU T, DING X L, LI J F. Review on *K* shortest paths algorithms[J]. Computer Engineering and Design, 2013, 34(11): 3900-3906, 3911.

[10] 王奇超, 文再文, 蓝光辉, 等. 优化算法的复杂度分析[J]. 中国科学: 数学, 2020, 50(9): 1271-1336.
WANG Q C, WEN Z W, LAN G H, et al. Complexity analysis for optimization methods[J]. Scientia Sinica (Mathematica), 2020, 50(9): 1271-1336.

[作者简介]

陆龑（1985− ），女，中国联合网络通信集团有限公司研究院工程师，主要研究方向为网络与信息安全。

徐雷（1978− ），男，博士，中国联合网络通信集团有限公司研究院教授级高级工程师，主要研究方向为网络与信息安全。



张曼君（1980− ），女，博士，中国联合网络通信集团有限公司研究院高级工程师，主要研究方向为网络与信息安全。