

5G 网络切片中的虚拟网络功能管理器映射算法

徐一凡^{1 2} 孟旭东^{1 2}

(1. 南京邮电大学, 江苏 南京 210003;

2. 江苏省电信网络融合实验室, 江苏 南京 210003)

摘 要: 5G 网络通过基于网络功能虚拟化(NFV)的切片技术来实现对不同应用场景下的网络需求,而切片中的不同类型的网络服务和虚拟网络功能分别由 NFV 的管理与编排(MANO)框架中的网络功能虚拟化编排器(NFVO)和虚拟网络功能管理器(VNFM)负责管理。针对在切片的实际部署中,NFVO 和 VNFM 因为延迟以及 VNF 负载变化而产生的性能问题,建立了一个基于成本的整数线性规划(ILP)最优化模型,并提出了一种基于禁忌搜索的 VNFM 映射算法,旨在符合切片性能要求的同时最大限度地降低切片的运营成本。该算法利用禁忌算法能够通过灵活地记忆避免迂回搜索,而快速找到全局最优解的特性,找出在切片中 VNFM 的最佳映射策略。实验结果表明,该算法能够明显提高切片网络资源利用率和降低切片网络的运营成本。

关键词: 5G; 网络切片; NFV 管理和编排; 虚拟网络功能管理器映射; 禁忌搜索

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2019)03-0041-06

doi: 10.3969/j.issn.1673-629X.2019.03.008

Virtual Network Function Manager Mapping Algorithm in 5G Network Slice

XU Yi-fan^{1 2}, MENG Xu-dong^{1 2}

(1. Nanjing University of Posts and Telecommunications, Nanjing 210003, China;

2. Jiangsu Telecom Network Fusion Laboratory, Nanjing 210003, China)

Abstract: 5G network achieves network requirements for different application scenarios through slicing technology based on NFV, and different types of network services and virtual network functions in the slice are managed by the NFVO and VNFM of the MANO framework of NFV. In the actual deployment for slicing, NFVO and VNFM may have performance issues due to some changes of network latency and VNF load. To solve this issue, we establish a cost-based ILP optimization model, and propose a VNFM embedding algorithm based on tabu search, which is to minimize slice operating costs while meeting slice performance requirements. As Tabu search, this algorithm can avoid duplicate search to find the global optimal solution by flexible memory. It can also find out the best embedding strategy for VNFM in slices. Experiment shows that the algorithm can significantly improve the utilization of slice network resources and reduce the operating costs of the slice network.

Key words: 5G; network slicing; NFV MANO; VNFM embedding; Tabu search

1 概述

5G 网络中通过虚拟化将一个物理网络分成多个切片网络来应对不同的应用场景^[1]。网络切片是一组虚拟网络功能(virtualized network function, VNF)及其资源组成的一个完整的逻辑网络^[2]。5G 中使用 NFV 的管理和编排(management and orchestration, MANO)模块来实现切片的部署和管理。

在欧洲电信标准协会(European telecommunication standards institute, ETSI)定义的 NFV 架构框架中,包括 VNF、网络功能虚拟化基础设施(network function virtualization infrastructure, NFVI)和 MANO 框架^[3]。VNF 是网络功能的软件实现,NFVI 是 VNF 部署的环境,是硬件和软件资源的组合,可以跨越分散的地理位置,其节点称为 NFVI 存在点(NFVI-PoP)。

收稿日期: 2018-05-05

修回日期: 2018-09-12

网络出版时间: 2018-12-20

基金项目: 国家科技重大专项(2012ZX03001008-003, 2011ZX03005-004-03); 国家“973”重点基础研究发展计划项目(2013CB329005)

作者简介: 徐一凡(1993-),男,硕士研究生,CCF 会员(92298G),研究方向为 5G 网络虚拟化、5G 切片技术;孟旭东,副教授,研究方向为电信网络和 IP 网络的交换、异构网络集成及业务融合、未来互联网体系结构、网络计算与分布式处理。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20181219.1553.084.html>

NFV MANO 框架能够在满足运营商要求(例如性能和可靠性)的同时,实现资源、VNF 和网络服务(network service, NS)的自动化管理^[4]。它包含三个功能块:虚拟化基础设施管理器(virtualized infrastructure manager, VIM)、VNFM 和 NFVO。MANO 模块通过一组已定义好的参考点与 NFV 架构框架中的其他功能模块进行通信^[5]。

在 MANO 中, VIM 负责对整个 NFVI 资源的管理和监控,例如管理配置硬件资源和虚拟化资源、收集和报告资源故障和性能信息等。VNFM 负责 VNF 的生命周期管理,包括 VNF 的实例化、VNF 容量的缩放、VNF 更新和改进、VNF 的终结(释放 VNF 占用的 NFVI 资源)。NFVO 负责在操作域内 NS 的管理和相应策略的制定,并且通过与 VNFM 的协作确保 VNF 实例满足性能与可靠性要求^[6]。

鉴于 MANO 功能模块所涵盖的角色,它在物理网络中的部署位置对切片的性能有重要影响。因为 MANO 功能块之间通过广域网(wide area network, WAN)链路进行通信,所以在实际的切片网络中,延迟是难以避免的。这些延迟因地点不同而有所不同,并随着网络流量的改变而改变^[7]。MANO 的管理功能无法承受高延迟,例如,延迟会阻止频繁收集和分析来自 VNF 实例和 VIM 的监控数据^[8]。另外,对于 VNF 故障管理来说,需要快速的故障通知和恢复,以尽量减少故障的影响并保持切片中网络服务的可靠性^[9]。所以说降低延迟对于整个切片的的管理的影响非常重要。

另外, MANO 模块的部署位置也会影响整个切片的运营成本,当系统中有很多的 VNFM 时这一点尤为明显。实际上,资源成本(包括计算和带宽资源成本)因网络运营商的不同、地点和时间的不同而有所差异,并且网络流量在系统中动态变化^[10]。因此,相应地调整 VNFM 的数量和位置可以为切片的运营节约大量的成本。所以 MANO 功能块的映射确实是一个需要解决的重要问题。

因为 VIM 的部署问题是 NFVI 设计的一部分,所以暂且不考虑。另外,忽略 NFVO 以缩小问题的范围。所以文中主要讨论的是切片系统中 VNFM 的位置映射问题。VNFM 的位置问题可以映射到虚拟网络映射问题(virtual network embedding, VNE)来解决^[11]。VNE 问题是 NP-Hard 问题,目前在研究中是通过启发式的算法(例如遗传算法)来解决这个问题^[12-13]。

文中旨在找到 VNFM 的最佳数量和位置,以在每个时刻都可以在延迟和性能限制下最大限度地降低运营成本。文中的主要贡献包括:

(1) 建立了 VNFM 映射(VNFM embedding, VN-

FME) 数学模型。主要侧重于动态 VNFME(即随着映射策略随着时间的推移而变化)的研究,并提出了此问题的 ILP 公式。

(2) 提出了一种基于禁忌搜索的映射算法来解决 VNFME 问题。禁忌搜索是一种使用自适应内存的高效邻域搜索方法。根据该问题的特殊性,详细设计了算法步骤。

2 VNFM 映射模型

2.1 问题描述

5G 网络切片通过网络资源与部署位置解耦、切片内资源动态伸缩调整,不仅可以提高 5G 网络服务的灵活性和资源利用率,而且降低了 5G 网络的部署与运营成本^[14]。也就是说,在切片中,网络服务和 VNF 可以在保证效率成本的同时,随时随地按需实例化,并进行弹性的容量缩放以满足不断变化的网络需求。因此,切片中 VNF 实例的数量、类型(如防火墙)和位置可能随时间变化。而对于 VNFM 来说,在任何时候,它的数量都应该适应部署在系统中的 VNF 实例,并保证其管理功能的可靠性,以实现切片系统的最佳性能。

由于 VNF 实例部署在跨地理分布的 NFVI-PoP 上,因此 NFVI-PoP 网络成为影响切片性能和运营成本的重要因素。如图 1, VNFM 通过 $Ve-Vnfm-vnf$, $Ve-Vnfm-em$, $Vi-Vnfm$ 和 $Or-Vnfm$ 等参考点与 VNF 实例、网元管理器(element manager, EM)、VIM 和 NFVO 交互。另一方面,实际中 VNFM 与这些功能块之间通过 WAN 链路进行通信,故 VNFM 在切片中的映射在确定各个参考点的延迟方面起着至关重要的作用。因此,计划外的 VNFM 位置会导致计划外的延迟,从而对系统的性能和可靠性产生负面影响。为了保证性能,每个参考点可以受到延迟限制的约束。

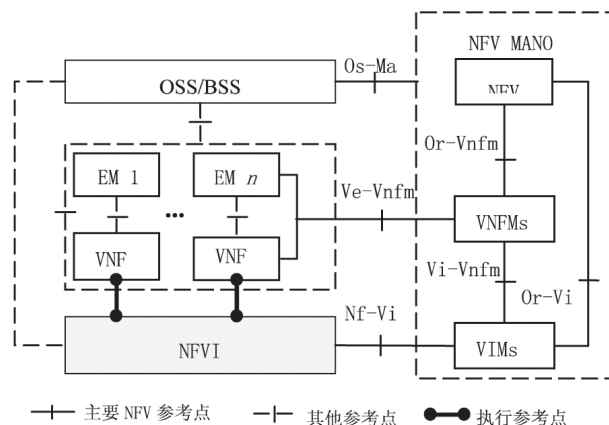


图 1 ETSI NFV 框架架构

因此, VNFM 的位置和数量对整个切片的性能以及成本至关重要。文中将相关问题称为 VNFM 映射(VNFME)问题。VNFME 问题具体定义如下: 给定

VNF实例的位置和NFVO的位置,目标是找到:(1)管理VNF实例所需的最佳VNFM数量;(2)VNFM的类型(例如通用VNFM,用于管理来自特定VNF的VNFM等);(3)VNFM在NFVI-PoP上的最佳位置。文中旨在满足切片系统中的通信延迟和容量限制(VNFM的容量限制)的同时,以最低的运营成本实现此目标。

VNFME分为静态和动态。在静态VNME中,映射和与VNF实例的关联是永久性的,不会随着时间而改变,这适用于系统中的更改(例如VNF实例数量)对于重新调整VNFM放置无关紧要的情况。但是,在实际的切片系统中,由于资源变化情况复杂,静态的映射策略并不能很好地适应这些变化,常常导致VNFM的数量和位置的错误配置。相比之下,动态VNFME旨在使VNFM数量和映射更好地适应变化。

2.2 系统模型

模型通过一组快照进行操作,定义一个快照 t 作为固定时间间隔内系统状态的表示。系统状态(如VNF实例的数量,系统中的延迟)可能因快照而异。因此,当系统从一个快照转移到下一个时,该模型考虑四种机制来适应系统以响应变化:添加新的VNFM以应付VNF实例的增加;删除现有的VNFM降低成本;将现有的VNFM迁移到新的位置;将VNF实例重新分配给另一个VNFM。

使用无向结构图 $G=(P,E)$ 来表示切片网络。其中 P 表示NFVI-PoP的节点 p 的集合, E 是连接这些节点的边的集合。边 $(p,q) \in E$ 连接了一对NFVI-PoP节点 p 和 q ,表示两点之间逻辑上的连通。快照 t 中 $\gamma_{p,q}(t)$ 和 $\delta_{p,q}(t)$ 表示边 $(p,q) \in E$ 的容量和延迟, $c_p^{\text{com}}(t)$ 和 $c_{p,q}^{\text{net}}(t)$ 表示 $p \in P$ 上计算资源成本和边 $(p,q) \in E$ 上的网络带宽成本。

假设NFVO部署在给定的NFVI-PoP上。用 $h_p \in \{0,1\}$ 表示它的位置,如果NFVO被放置在 $p \in P$,则 h_p 等于1,否则等于0。

$M(t)$ 表示可用于快照 t 的VNFM的集合。 K 表示VNFM或VNF类型的集合。 n_k 表示一个类型为 $k \in K$ 的VNFM m 的容量,即VNFM可以管理的VNF实例的最大数量。 $M_k(t) \in M_k$ 表示可用于快照 t 上的类型为 $k \in K$ 的VNFM的集合。 $\hat{M}_k(t)$ 表示从 $M_k(t)$ 中被选择用于快照 t 的类型为 $k \in K$ 的VNFM的集合,即活动的VNFM。 $g_{m,k}$ 指代迁移类型为 $k \in K$ 的VNFM在之前的NFVI-PoP和新的NFVI-PoP之间的边上所消耗的带宽。 $c_{m,k}^{\text{mig}}(t)$ 表示迁移类型为 $k \in K$ 的VNFM的成本,也代表了由迁移造成的服务中断的惩罚。

$V(t)$ 表示在快照 t 上管理的VNF v 实例的集合。 $V_k(t)$ 是 $V(t)$ 一个子集,其中包括需要被 $k \in K$ 类型

VNFM管理的所有VNF实例。 $l_{v,p}(t) \in \{0,1\}$ 表示VNF实例的位置,如果VNF实例在快照 t 上被放置在 $p \in P$,则 $l_{v,p}(t)$ 等于1,否则为0。此外,通过两个延迟限制来控制生命周期管理中的通信开销。第一个限制是 $\varphi_v(t)$,表示VNF实例 v 与通过快照 t 管理它的VNFM之间的最大允许延迟 v 。第二个限制是 $\omega_v(t)$,是指NFVO与快照 t 上VNFM之间的最大允许延迟。 S_v 和 \hat{S}_v 则表示在NFVO与旧VNFM之间以及NFVO与新VNFM之间重新分配VNF实例 v 时消耗的带宽。并且 $c_v^{\text{rea}}(t)$ 表示重新分配的成本,也代表了重新配置系统以确保其稳定性所付的代价。

MANO功能模块与其他非MANO功能模块(如EM)交互以管理VNF实例的生命周期。对于VNF实例 v ,假设 $u_v^{O,M}(t)$, $u_v^{O,I}(t)$, $u_v^{M,I}(t)$, $u_v^{M,V}(t)$ 分别表示快照 t 上在参考点Or-Vnfm,Or-Vi,Vi-Vnfm和Ve-Vnfm上,NFVO和VNFM,NFVO和VIM,VNFM和VIM,VNFM和EM/VNF实例 v 之间的通信期间消耗的带宽。

2.3 问题表述

将VNFME制定为ILP问题,目标是根据单个快照制定映射策略,以最小化切片的运营成本。对于静态VNFME,只在一个快照 t 上操作,代表系统的永久状态,从系统的永久配置中推导公式。而对于动态VNFME,是对系统中的每对连续快照 $t-1$ 和 t 进行操作。更确切地说,给定快照 $t-1$ 和 t ,在快照 $t-1$ 结束时,决定VNFM的位置以及它们与VNF实例 v 在快照 t 上关联。所有在 $M(t)$ 中的VNFM的放置决定是否添加新的VNFM,以及保留,删除或迁移现有的VNFM。 $M_k(t)$ 定义如下:

$$M_k(t) = F_k(t) \cup \hat{M}_k(t-1) \quad (1)$$

其中, $F_k(t)$ 是一系列可以被添加到系统中的快照 t 上类型为 k 的新VNFM的集合,比如:

$$|F_k(t)| = \sum_{p \in P} \left(\frac{\sum_{v \in V_k(t)} l_{v,p}(t)}{n_k} \right) \quad (2)$$

其中, $|F_k(t)|$ 表示可以在快照 t 中被添加进系统的 k 类型的VNFM数量的上限, $\hat{M}_k(t-1)$ 是快照 $t-1$ 活动的VNFM的集合。决策变量为 $x_{m,k,p}(t)$ 、 $y_{v,m,k,p}(t)$ 。当 $m \in M_k(t)$ 映射在 $p \in P$ 上时, $x_{m,k,p}(t)$ 的值为1,否则为0;当 $v \in V_k(t)$ 在 $p \in P$ 上分配给 $m \in M_k(t)$ 进行管理时, $y_{v,m,k,p}(t)$ 的值为1,否则为0。

运营成本的定义分为四个不同的部分,定义如下:

生命周期管理成本($C^{\text{lif}}(t)$):表示快照 t 上系统执行所有VNF实例的生命周期管理所消耗的网络通信带宽的成本:

$$C^{\text{lif}}(t) = \sum_{(p,q) \in E} [B^{\text{lif}}(p,q) + B^{\text{lif}}(q,p)] c_{p,q}^{\text{net}}(t) \quad (3)$$

其中

$$B^{\text{lif}}(p, q) = \sum_{k \in K} \sum_{v \in V_k(t)} \sum_{m \in M_k(t)} \{y_{v, m, k, p}(t) h_q u_v^O(t) + y_{v, m, k, p}(t) l_{v, q}(t) [u_v^{M, J}(t) + u_v^{M, J}(t)] + l_{v, p}(t) h_q u_v^O(t)\}$$

计算资源成本($C^{\text{com}}(t)$): 表示在快照 t 上分配给 VNFM 的计算资源成本:

$$C^{\text{com}}(t) = \sum_{k \in K} \sum_{m \in M_k(t)} \sum_{p \in P} x_{m, k, p}(t) c_p^{\text{com}}(t) \quad (4)$$

在此假设 VNFM 需要单个计算资源单元。这个假设是由于缺乏关于 MANO 功能模块的资源分配的可用数据。

迁移成本($C^{\text{mig}}(t)$): 表示在从快照 $t-1$ 转换到快照 t 的同时, 将 VNFM 从一个 NFVI-PoP 迁移到另一个 NFVI-PoP 时所隐含的成本。

$$C^{\text{mig}}(t) = \sum_{k \in K} \sum_{m \in M_k(t-1)} \sum_{p, q \in E} [x_{m, k, p}(t) x_{m, k, p}(t-1) + x_{m, k, p}(t) x_{m, k, p}(t-1)] c_{m, k}^{\text{mig}}(t) \quad (5)$$

重新分配成本($C^{\text{rea}}(t)$): 当从快照 $t-1$ 转换到快照 t 时, 系统中剩余的 VNF 实例可能会重新分配给新的 VNFM。按如下方式计算重新分配这些 VNF 实例的成本:

$$C^{\text{rea}}(t) = \sum_{k \in K} \sum_{v \in V_k(t) \cap V_k(t-1)} \sum_{m \in M_k(t)} \sum_{p \in P} y_{v, m, k, p}(t) (1 - \sum_{q \in P} z_{v, m, k, q}(t)) c_v^{\text{rea}}(t) \quad (6)$$

其中, 当 $v \in V_k(t) \cap V_k(t-1)$ 并且 $m \in M_k(t-1)$ 时, $z_{v, m, k, p}(t) = y_{v, m, k, p}(t-1)$, 否则为 0。

最后, 优化问题的目标是最小化以上四个成本的加权和, 表示如下:

$$\text{Min } C^{\text{lif}}(t) + C^{\text{com}}(t) + C^{\text{rea}}(t) + C^{\text{mig}}(t) \quad (7)$$

在静态情况下, 式 7 不包括 $C^{\text{rea}}(t)$ 和 $C^{\text{mig}}(t)$ 。

约束条件如下:

(a) 每个 VNF 实例应该分配给一个 VNFM, 如式 8 所示:

$$\sum_{m \in M_k(t)} \sum_{p \in P} y_{v, m, k, p}(t) = 1, \forall k \in K, p \in V_k(t) \quad (8)$$

(b) VNF 实例分配给处于同一个 NFVI-PoP 上的 VNFM, 如式 9 所示:

$$y_{v, m, k, p}(t) \leq x_{m, k, p}(t), \quad \forall k \in K, p \in V_k(t), m \in M_k(t), p \in P \quad (9)$$

(c) 确保分配给每个 VNFM 的 VNF 实例数量不超过其容量, 如式 10 所示:

$$\sum_{v \in V_k(t)} y_{v, m, k, p}(t) \leq n_k, \quad \forall k \in K, p \in V_k(t), m \in M_k(t), p \in P \quad (10)$$

(d) VNFM 只能位于一个 NFVI-PoP。这个约束由式 11 定义:

$$\sum_{p \in P} x_{m, k, p}(t) \leq 1, \forall k \in K, m \in M_k(t) \quad (11)$$

(e) 确保 VNFM 仅在管理至少一个 VNF 实例时

处于活动状态, 如式 12 所示:

$$x_{m, k, p}(t) \leq \sum_{v \in V_k(t)} y_{v, m, k, p}(t), \quad \forall k \in K, p \in V_k(t), m \in M_k(t), p \in P \quad (12)$$

(f) 每个 VNF 实例有两个延迟限制来控制其分配的 VNFM 的参考点上的延迟。通过式 13 和式 14 来执行这些约束:

$$(y_{v, m, k, p}(t) l_{v, q} + y_{v, m, k, p}(t) l_{v, p}) \delta_{p, q}(t) \leq \varphi_v(t), \quad \forall k \in K, p \in V_k(t), m \in M_k(t), (p, q) \in E \quad (13)$$

$$(y_{v, m, k, p}(t) h_q + y_{v, m, k, p}(t) h_p) \delta_{p, q}(t) \leq \omega_v(t), \quad \forall k \in K, p \in V_k(t), m \in M_k(t), (p, q) \in E \quad (14)$$

(g) 式 15 保证每条边上的使用带宽不超过其容量:

$$B^{\text{lif}}(p, q) + B^{\text{lif}}(q, p) + B^{\text{mig}}(p, q) + B^{\text{mig}}(q, p) + B^{\text{rea}}(p, q) + B^{\text{rea}}(q, p) \leq \gamma_{p, q}(t), \quad \forall (p, q) \in E \quad (15)$$

其中

$$B^{\text{mig}}(p, q) = \sum_{k \in K} \sum_{m \in M_k(t-1)} x_{m, k, p}(t) x_{m, k, p}(t-1) g_{m, k} \\ B^{\text{rea}}(p, q) = \sum_{k \in K} \sum_{v \in V_k(t) \cap V_k(t-1)} \sum_{m \in M_k(t)} \{z_{v, m, k, p}(t) h_q (1 - \sum_{p \in P} y_{v, m, k, p}(t)) \hat{S}_v + y_{v, m, k, p}(t) h_q (1 - \sum_{p \in P} z_{v, m, k, p}(t)) S_v\}$$

在静态情况下, 式 15 不包括 $B^{\text{rea}}(t)$ 和 $B^{\text{mig}}(t)$ 。

3 算法描述

禁忌搜索是一种元启发式算法, 用于指导局部搜索过程以找到组合优化问题的近似最优解。它从初始解开始探索搜索空间, 并迭代地执行移动以从当前解决方案转移到其邻域中的另一个解决方案, 直到满足终止标准^[15]。文中设计的算法步骤如图 2 所示。

Algorithm 1 VNF embedding algorithm based on tabu search

```

1:  $S_0(t) \leftarrow \text{GreedyInitialSolution}$ 
2:  $S_{\text{current}}(t) \leftarrow S_0(t), S_{\text{best}}(t) \leftarrow S_0(t), j \leftarrow 0$ 
3: repeat
4: moves-list  $\leftarrow$  Create Candidate moves list
5: best-move  $\leftarrow$  Choose from move-list the move
6: that generates the best solution
7: apply best-move on  $S_{\text{current}}(t)$  solution
8: add best-move to tabu list for  $i_{\text{tabu}}$  iterations
9:  $j \leftarrow j + 1$ 
10: if  $f(S_{\text{current}}(t)) < f(S_{\text{best}}(t))$  then
11:  $S_{\text{best}}(t) \leftarrow S_{\text{current}}(t)$ 
12:  $j \leftarrow 0$ 
13: end if
14: until  $j = i_{\text{stop}}$ 
15: return  $S_{\text{best}}(t)$ 

```

图 2 基于禁忌搜索的 VNFM 映射算法伪代码

3.1 初始映射策略

快照 t 的初始映射策略 $S_0(t)$ 可以使用贪婪算法来求解。主要有两个步骤, 第一步在系统中存在快照

$t-1$ 时开始执行。在这一步中, 实现了根据快照 $t-1$ 作出决策。因此, 将每个活动的 VNFM $m \in \hat{M}_k(t)$, $\forall k \in K$ 映射在同一个 NFVI-PoP 并将每个 $v \in V_k(t) \cap V_k(t-1)$, $\forall k \in K$ 分配给同一个 VNFM。确定最佳机制(例如, VNFM 迁移和 VNF 重新分配)以重新调整当前系统配置非常具有挑战性。因此, 这一步骤的目标是维持初始解决方案中的当前配置, 并将决策责任委托给禁忌搜索算法, 该算法可以评估所有可能的替代方案并作出决策。

在第二步中, 将每个新的 VNF 实例 $v \in V_k(t) \setminus V_k(t-1)$, $\forall k \in K$ 分配到具有足够容量且位于与 VNF 实例相同的 NFVI-PoP 的活动 VNFM。将 VNFM 及其分配的 VNF 实例置于相同位置的目的是为了保持初始映射策略中的总延迟最小化并随后使禁忌搜索迭代最小化。如果没有活动的 VNFM 满足上述条件, 新的 VNFM $m \in M_k(t) \setminus \hat{M}_k(t)$, $\forall k \in K$ 将被激活并放置在 VNF 实例所在的 NFVI-PoP 处, 然后这些 VNF 实例分配给新的 VNFM 来管理。

此外, 对于新的 VNFM $m \in M_k(t) \setminus \hat{M}_k(t)$, $\forall k \in K$, 如果在初始映射策略中没有被分配 VNF 实例, 则被视为被禁用的 VNFM, 在后面将由 Tabu 搜索算法来激活。生成的初始解决方案满足所有约束条件。

3.2 邻域移动

文中的禁忌搜索算法采用四种移动类型来生成邻域解, 定义如下:

VNF 重新分配: 随机选择一个 VNF 实例将其重新分配给另一个 VNFM。新的 VNFM 可能处于活动状态或不活动状态。如果是不活动的, 则 VNFM 被激活。如果旧 VNFM 不再有与其关联的 VNF 实例, 则停用。

VNFM 重新定位: 随机选择活动的 VNFM 并移动到另一个 NFVI-PoP。选择新的位置, 以便不违反所有分配的 VNF 实例的延迟限制。

批量 VNF 重新分配: 随机抽取活跃的 VNFM。然后, 如果剩余的活动 VNFM 具有足够的容量来管理其分配的 VNF 实例, 同时满足它们的延迟约束, 则重新分配 VNF 实例, 并且停用所选 VNFM。

VNFM 取消激活: 让 $P_m(t) \subseteq P$ 表示一个活动的 VNFM $m \in \hat{M}_k(t)$, 可以在满足其对应的 VNF 实例的延迟约束的情况下映射于此。然后, 如果两个 VNFM m 和 \hat{m} 有重叠覆盖(即 $\{P_{m,\hat{m}}(t) = P_m(t) \cap P_{\hat{m}}(t)\} \neq \emptyset$), 则需要随机选择两个 VNFM m 和 \hat{m} 并评估重叠覆盖率。如果评估达标并且 VNFM m 有足够的容量为分配给 \hat{m} 的 VNF 实例提供服务, 那么将这些 VNF 实例重新分配给 m , \hat{m} 则被停用。如果 VNFM m 尚未放置在 $p \in P_{m,\hat{m}}(t)$, 那么它被重新定位到 $P_{m,\hat{m}}(t)$ 。

3.3 禁忌列表和渴望水平条件

禁忌搜索使用称为禁忌表的内存结构来记录有关最近搜索历史的信息, 这可以避免局部最优解, 并防止对先前得到的最优解的重复访问。在后面的迭代中, 如果移动已经存在于禁忌表中则被禁止。此外, 如果禁忌表中的移动达到渴望水平条件, 则被解禁并可以重新选择。

3.4 评价函数

在每次迭代中, 禁忌搜索算法评估一组候选移动并选择生成最优邻域解的移动。该算法使用分层目标函数(f)来评估邻域解, 其中主要目标首先被最小化, 然后对于相同的主要目标值, 次要目标被最小化。主要目标被定义为由模型目标函数和与解相关的总惩罚之和。次要目标是活动的 VNFM 与其相关的 VNF 实例之间的延迟总和。最终目标是最大化 $P_m(t)$ 对于所有活动的 VNFM, 然后检测并消除重叠覆盖。

3.5 终止准则

当找到的最佳解决方案对于一定数量的连续迭代 i_{stop} 次没有改进时, 则算法停止。 i_{stop} 被定义为 $25 \sqrt{|V(t)|} + |M(t)|$ 。此公式旨在允许 $V(t)$ 和 $M(t)$ 相对增长。

4 仿真实验与分析

4.1 仿真环境

仿真实验使用配置为 12 GB 内存, 64 位 Win10 操作系统, Intel Core i5 处理器的计算机进行评估, 使用 MATLAB 进行编程, 实验中的物理网络及虚拟网络拓扑都采用 GT-ITM 生成。

预设参数信息如下:

每个 NFVI-PoP 计算资源的成本($c_p^{\text{com}}(t)$)服从 $[0.5, 1]$ 均匀分布。边缘相关的带宽成本($c_{p,q}^{\text{net}}(t)$)设置为 0.8 元/GB。所有边缘的容量($\gamma_{p,q}(t)$)设置为 10 Gb/s。迁移 VNFM 所消耗的带宽($g_{m,k}$)设置为 40 GB。迁移成本($c_{m,k}^{\text{mig}}(t)$)等于在迁移中消耗的带宽的成本 40×0.8 元。VNFM 容量(n_k)设置为 6。

VNF 实例有两种类型 C_1 和 C_2 。 C_1 包含复杂的高吞吐量事务 VNF, 如服务呼叫会话控制功能, 需要严格的可靠性和性能要求。 C_2 包括简单的 VNF, 例如可用于住宅虚拟客户端设备(vCPE)的防火墙。在这种情况下, VNF 性能异常会影响较小。为了简单起见, 假设 S_v 和 \hat{S}_v (在 VNF 实例重新分配中消耗的带宽)相等为 2 MB。VNF 再分配成本 $c_v^{\text{rea}}(t)$ 是重新分配中消耗的带宽的成本, 等于 $\frac{(2+2) \times 0.8}{1024}$ [¥]。仿真中

关于网络流量和延迟的参数设置如表 1 所示。

表 1 仿真参数

参数	C_1 取值	C_2 取值
VNFM 和 VNF 实例、EM、VIM 之间的最大延迟 $\varphi_v(t)$ /ms	35	60
NFVO 和 VNFM 之间的最大延迟 $\omega_v(t)$ /ms	70	100
NFVO 和 VNFM 之间消耗的带宽 $u_v^O(t)$ /(MB/t)	0.51	0.11
NFVO 和 VIM 之间消耗的带宽 $u_v^O(t)$ /(MB/t)	0	0
VNFM 和 VIM 之间消耗的带宽 $u_v^M(t)$ /(MB/t)	1.26	0.33
VNFM 和 VNF/EM 之间消耗的带宽 $u_v^M(t)$ /(MB/t)	1.77	0.31

4.2 结果分析

文中将从成本和 VNFM 数量这两个方面对基于禁忌搜索的 VNFM 映射算法的性能进行评估。为了更加直观地了解此算法的性能,将使用此算法获得的结果与基于贪婪算法获取的初始结果和基于遗传算法的映射算法取得的结果进行比较。图 3 为三种算法在不同的 VNF 实例数量的情况下,其最优映射策略下切片运营总成本的对比。可以看出,提出的基于禁忌搜索的 VNFM 映射算法总能在初始解的基础上产生更优的解,一定程度上降低了切片的运营成本,验证了该算法的有效性。同时可以看出这些解与基于遗传算法产生的解相比基本一致,有时甚至更加优秀。

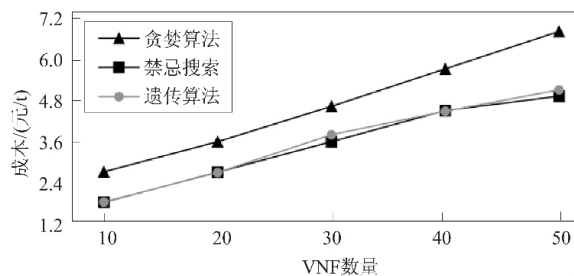


图 3 三种算法总成本对比

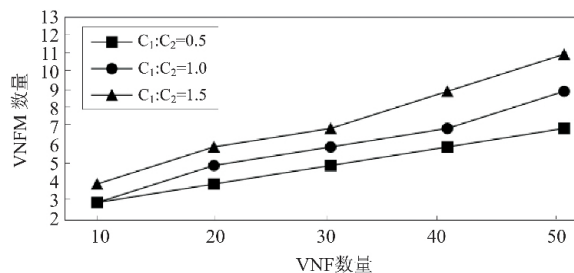
图 4 C_1 和 C_2 不同比例对所需 VNFM 数量的影响

图 4 描述的是 VNF 实例的数量和 C_1, C_2 在 VNF 总量中所占不同比重对切片系统中所需的 VNFM 数量的影响。可以看出,随着 VNF 数量的增加,VNFM 容量有限,所以 VNFM 数量必然也需要增加以管理更多的 VNF,但是随着 C_1 类型的 VNF 所占比重的增加,系统所需 VNFM 的数量更多,并且其增长速率也更高。这是因为 C_1 类型的 VNF 的性能和延迟要求较高,所以对管理它的 VNFM 的性能要求也就相应的高。要想保证切片系统的性能,系统只能使用相匹配的高性能的专用 VNFM 来管理,或者减少通用型 VN-

FM 的容量来保证其管理性能的可靠性,所以相比于性能要求不高的 C_2 类型的 VNF, C_1 类型的 VNF 需要更多的 VNFM 来管理。

图 5 为不同快照中随着系统中 VNF 实例数量的变化,动态 VNFM 与静态 VNFM 产生的切片运营总成本对比。可以看出,动态 VNFM 产生的成本要低于静态情况,并且随着 VNF 数量的变化,动态 VNFM 下成本的变化相对对应较为平缓。而静态的 VNFM 显然无法很好地适应切片系统状态的变化,并且成本也更高。

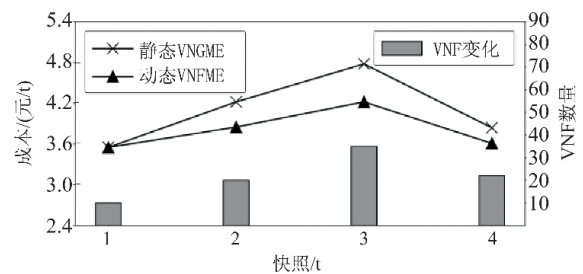


图 5 动态 VNFM 与静态 VNFM 总成本对比

5 结束语

介绍并研究了 5G 网络切片中的 VNFM 问题。通过分析实际网络切片系统中的延迟与成本因素,建立了问题的整数线性规划最优化模型,并提出了禁忌搜索算法来获得最优的映射策略。实验结果表明,禁忌搜索算法总能得到使切片运营成本最小的映射策略。并且在动态 VNFM 中,相对于静态 VNFM,该算法可以使策略适应系统中的变化,从而使切片运营成本显著降低。

参考文献:

- [1] 任 驰,马瑞涛.网络切片网络切片:构建可定制化的 5G 网络[J].中兴通讯技术 2018 24(1):26-30
- [2] 月 球,肖子玉,杨小乐.未来 5G 网络切片技术关键问题分析[J].电信工程技术与标准化 2017 30(5):45-50.
- [3] MA Shicong, WANG Baosheng, ZHANG Xiaozhe et al. ApplianceBricks: a scalable network appliance architecture for network functions virtualization[J]. China Communications, (下转第 50 页)



(2) 识别成功

图 2 实时识别过程

6 结束语

提出一种 Gabor 和 LBPH 结合的实时人脸识别算法,识别率高,相应系统具有在线管理方式,节省了人工成本。在 Exynos 4412 嵌入式平台所开发的实时人脸识别系统,充分利用了硬件系统的性能。搭建的小型化人脸识别系统为人工智能、物联网、考勤系统等提供了一种解决方案,有望推广到各类身份识别的应用场景。

参考文献:

- [1] 李 健. 基于 Gabor 变换和 LBP 相结合的人脸识别算法的研究[D]. 太原: 太原理工大学, 2016.
- [2] 吴 帅. 基于 Gabor 滤波器的伸长局部二值模式算法在视频人脸识别系统中的应用研究[D]. 昆明: 昆明理工大学, 2016.
- [3] TURK M, PENTLAND A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.
- [4] KIRBY M, SIROVICH L. Application of the Karhunen-Loeve procedure for the characterization of human faces[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 12(1): 103-108.
- [5] 岑 瑶, 潘 新, 郝晓晶, 等. 基于 MB-LBP 和 HOG 的掌纹识别[J]. 计算机应用研究, 2017, 34(3): 920-923.
- [6] 徐 丹, 于 龙, 陈唐龙, 等. 基于 LBP-HOG 特征的接触网旋转双耳区域识别[J]. 铁道科学与工程学报, 2017, 14(2): 370-378.
- [7] 吴凌智, 陈 玮. 基于 MBLBP 和 Gabor 小波的掌纹识别[J]. 软件导刊, 2017, 16(2): 19-22.
- [8] HWANG W, HUANG X, LI S Z, et al. Face recognition using extended curvature Gabor classifier bunch[J]. Pattern Recognition, 2015, 48: 1247-1260.
- [9] LIU Jun, JING Xiaojun, SUN Songlin, et al. Local Gabor dominant direction pattern for face recognition[J]. Chinese Journal of Electronics, 2015, 24(2): 245-250.
- [10] 张 露, 屈诗云, 庄 箫. 基于 Gabor 小波自适应判别的单样本人脸识别[J]. 工业控制计算机, 2015, 28(1): 95-96.
- [11] OJALA T, PIETIKÄINEN M, HARWOOD I. A comparative study of texture measures with classification based on feature distributions[J]. Pattern Recognition, 1996, 29(1): 51-59.
- [12] 徐向茹. 基于人脸识别身份验证系统的研究与实现[D]. 成都: 电子科技大学, 2016.
- [13] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. Kauai, HI, USA: IEEE, 2003: 511-518.
- [14] 刘 贝. 基于嵌入式的人脸识别智能门禁系统[D]. 西安: 西安工业大学, 2016.
- [15] ALAMEDDINE H A, SEBBAH S, ASSI C. On the interplay between network function mapping and scheduling in VNF-based networks: a column generation approach[J]. IEEE Transactions on Network and Service Management, 2017, 14(4): 860-874.
- [16] 邵维专, 吕光宏. 网络功能虚拟化资源配置及优化研究综述[J]. 计算机应用研究, 2018, 35(2): 321-326.
- [17] VASSILARAS S, GKATZIKIS L, LIAKOPOULOS N, et al. The algorithmic aspects of network slicing[J]. IEEE Communications Magazine, 2017, 55(8): 112-119.
- [18] 肖蔼玲, 王 颖, 孟洛明, 等. 基于知识描述和遗传算法的跨域虚拟网络映射[J]. 软件学报, 2014, 25(10): 2189-2205.
- [19] 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.
- [20] 钱 昭. 浅谈 5G 移动网络切片技术及关键问题研究[J]. 网络安全技术与应用, 2017(12): 99.
- [21] 程 航, 张 磊. 求解最短路问题的改进禁忌搜索算法[J]. 交通科技与经济, 2018, 20(2): 35-38.

(上接第 46 页)

2016, 13(1): 32-42.

- [4] CASELLAS R, MARTINEZ R, VILALTA R. Control, management, and orchestration of optical networks: evolution, trends and challenges view document[J]. Journal of Light-wave Technology, 2018, 36(7): 1390-1402.
- [5] MAMATAS L, CLAYMAN S, GALIS A. Information exchange management as a service for network function virtualization environments[J]. IEEE Transactions on Network and Service Management, 2016, 13(3): 564-577.
- [6] 王海宁, 赵慧玲. NFVO 标准和实践[J]. 电信科学, 2017, 33(4): 3-9.
- [7] LANGE S, GEBERT S, ZINNER T, et al. Heuristic approaches to the controller placement problem in large scale SDN networks[J]. IEEE Transactions on Network and Service Management, 2015, 12(1): 4-17.
- [8] MIJUMBI R, SERRAT J, GORRICO J, et al. Management and orchestration challenges in network functions virtualization[J]. IEEE Communications Magazine, 2016, 54(1): 98-105.