

# 基于 INT 感知的 5G 核心网用户面流量调度系统

王超, 任蒙, 王尚广

(北京邮电大学网络与交换技术全国重点实验室, 北京 100876)

**摘要:** 为解决 5G 移动数据流量激增和新型网络应用不断涌现给网络运维造成巨大压力的问题, 设计并实现一个面向 5G 核心网用户面的流量调度系统, 包含网络状态信息感知子系统和路由决策子系统。在网络状态信息感知子系统中, 传统带内网络遥测方法带宽开销高, 而且不是专门为无线网络系统设计的, 此外, 在 5G 核心网中应用带内网络遥测还面临测量精度不高、QoS 无法保障等挑战。为此, 提出了基于带内网络遥测的 5G 核心网用户面状态信息感知方案, 将遥测信息插入 GTP-U 报文的扩展头部, 实时检测 UPF 状态信息, 实现网络状态动态随路测量。路由决策子系统实现了基于改进蚁群算法的流量调度算法, 利用感知到的网络状态信息升级信息素函数更新方式, 完成基于实时网络状态的路由决策。实际部署测试结果表明, 网络状态信息感知子系统能正常感知网络信息, 而且所做路由决策在时延、吞吐量和丢包率方面优于传统路由算法。

**关键词:** 5G 核心网; 用户面; 流量调度; 带内网络遥测; 改进蚁群算法

**中图分类号:** T393

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2023208

## Traffic scheduling system of the 5G core network user plane based on INT perception

WANG Chao, REN Meng, WANG Shangguang

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

**Abstract:** To solve the problem of huge pressure on network operation caused by the surge of 5G mobile data traffic and the continuous emergence of new network applications, a traffic scheduling system of the 5G core network user plane was designed and implemented, including network state information perception subsystem and routing decision subsystem. In the network state information perception subsystem, traditional in-band network telemetry methods incurred high bandwidth overhead and were not specifically designed for wireless network systems. Furthermore, the application of in-band network telemetry in the 5G core network faced challenges such as low measurement accuracy and the inability to guarantee QoS. A 5G core network user plane state information perception scheme based on in-band network telemetry was proposed. The telemetry information was inserted into the extended header of GTP-U message realize the measurement of network state along the path. The routing decision subsystem implemented the traffic scheduling algorithm based on the improved ant colony algorithm. The update method of the pheromone function was promoted by the perceived status information to complete the routing decision based on the real-time network state. The deployment test results showed that the network state information perception subsystem can perceive network information normally, and the routing decisions were superior to traditional routing algorithms in terms of delay, throughput and packet loss rate.

**Keywords:** 5G core network, user plane, traffic scheduling, in-band network telemetry, improved ant colony algorithm

收稿日期: 2023-04-07; 修回日期: 2023-07-03

通信作者: 王尚广, sgwang@bupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62032003, No.U21B2016)

**Foundation Item:** The National Natural Science Foundation of China (No.62032003, No.U21B2016)

## 0 引言

在万物互联时代, 5G 有效支持了大量新型应用接入和大规模机器类型通信, 提供了毫秒级端到端通信时延、百万级连接密度和每秒数十吉比特的峰值速率, 用户服务体验明显提升<sup>[1]</sup>。随着工业 4.0 的到来, 元宇宙、虚拟现实、移动游戏等新型移动网络业务不断涌现, 移动通信数据流量急剧增长, 据国际电信联盟报告, 全球移动数据流量将在 2030 年达到每人每月 5 016 EB<sup>[2]</sup>。5G 核心网用户面负责业务数据流量的处理和转发, 如此庞大的流量负载会给 5G 网络运营及管理带来巨大压力。如果 5G 用户面不能及时路由数据流量, 引起的高时延将会给无人驾驶、智慧医疗等业务带来灾难性后果。因此, 对 5G 核心网用户面的流量调度机制进行改进与优化, 提升网络服务质量, 将对用户、服务提供商和基础设施提供商具有重要意义。

为避免差异化新兴业务需求与海量数据流量给 5G 网络造成网络拥塞和负载失衡等问题, 对业务流量进行合理规划与调度是最直接的解决方案。等价多路径 (ECMP, equal-cost multi-path) 算法<sup>[3]</sup>是广泛应用于数据中心网络 (DCN, data center networking) 的流量调度算法, 通过将不同流散列分配到多个等价路径上, 实现负载均衡。但是 ECMP 算法通过计算哈希值静态分配数据流路径, 容易将多数流分配到同一路径, 造成拥塞。Ma 等<sup>[4]</sup>提出一种面向 DCN 的基于蚁群算法的流量负载均衡调度算法, 该算法提升了链路利用率和网络吞吐量。Silva 等<sup>[5]</sup>提出一种用于软件定义网络 (SDN, software defined network) 数据平面的动态负载均衡算法, 通过动态调整流量路由路径, 明显降低时延和丢包率。对于 5G 核心网的流量调度, Tan 等<sup>[6]</sup>提出一种面向用户面的智能路由选择机制, 依据用户面流量负载选择开销最小的路由路径。此外, Nguyen 等<sup>[7]</sup>提出一种面向基于矢量包处理的 5G 核心网用户面设备的动态负载均衡算法, 考虑了矢量包处理向量速率和会话数量并将它们作为性能指标。但是该算法是基于矢量包处理平台实现的, 缺乏普适性。

虽然上述流量调度算法不同程度地实现了时延、吞吐量、链路利用率等性能优化, 但是在面向 5G 核心网的流量调度算法中, 只依靠某一时刻获取的用户面状态来制定流量调度决策, 不能跟随时变的用户面状态信息来调整路由决策, 缺乏连续性

和实时性。为了准确且实时获取 5G 核心网用户面网络状态信息, 使其作为路由决策依据, 需要引入网络测量技术<sup>[8]</sup>。

根据测量方式不同, 网络测量可以分为主动测量、被动测量和混合测量<sup>[9]</sup>。主动测量将探测包注入网络, 通过观察分析探测包状态, 进而推测网络性能, 如 Ping、Traceroute 等, 但是这种测量方式仅能测量网络时延, 局限性较大, 而且, 主动测量不能保证探测包与业务数据流经同一条路径, 引入了额外带宽开销, 因此测量结果往往不准确。被动测量在网络交换设备上利用流量镜像或端口转发来记录并分析数据流量, 如 NetFlow、sFlow、SNMP 等, 但是该测量技术受到交换设备性能限制, 测量精度较粗, 而且仅能测量本地设备状态, 缺乏对业务流量的端到端感知。混合测量则是尝试结合上述 2 种测量方式优点的一种网络测量技术, 带内网络遥测 (INT, in-band network telemetry) 是其典型代表。INT 将报文转发与网络测量结合, 通过在交换机、智能网卡等转发设备上收集设备状态元信息并将其插入报文, 从而实现随路测量。INT 因实时性强、测量精度高、测量数据类型丰富等优点引起了业界的广泛关注。

在基于 INT 的网络流量调度方面, Yao 等<sup>[10]</sup>提出一种 SDN 中的自学习策略控制架构, 为动态生成控制策略, 使用 INT 收集网络遥测信息, 依据遥测信息运行深度强化学习算法, 实现了网络的闭环控制。针对 IEEE 802.11 网络无法细粒度编程和管理多个网络的问题, Isolani 等<sup>[11]</sup>提出一种基于 INT 的 SDN 框架来收集细粒度网络统计信息, 之后基于 INT 测量的流量统计数据为应用需求进行网络切片, 以提高服务质量 (QoS, quality of service) 交付。Katta 等<sup>[12]</sup>提出一种基于虚拟机监视器的可扩展负载均衡器机制, 该机制采用路由追踪机制, 利用 INT 获取的链路拥塞情况或链路利用率来发现可用路径, 通过操纵报头字段, 将流量引导到可用路径上, 避免网络拥塞。

上述 INT 原型系统或方案实现主要基于 SDN, 以用户数据报协议 (UDP, user datagram protocol) 报文作为 INT 载体。INT 技术作为一种新型网络测量技术, 为网络拥塞控制、负载均衡等网络管理任务提供了细粒度和高精度的实时网络状态信息。但是, INT 最初是针对有线网络设计的, 将 INT 应用于 5G 核心网测量还是一个全新的尝试, 面临诸多

挑战。首先，INT 遥测报文携带遥测指令和遥测数据等固有信息，导致网络带宽开销较大；然后，5G 核心网数据传输速率和复杂性较高，对网络性能进行准确测量难度较大，需要保障 INT 技术测量结果的准确性；此外，将 INT 引入 5G 核心网还可能影响不同服务的 QoS。因此，本文拟研究基于 INT 的 5G 核心网用户面流量调度系统，以探究基于实时网络状态的流量调度增益。

5G 核心网采用基于服务的架构 (SBA, service-based architecture)，将控制面与用户面分离，用户面分布下沉到网络边缘，实现了用户面功能 (UPF, user plane function) 灵活部署，为用户就近接入和规划用户面路径提供了可能。为保证用户面流量通信的可靠性，可以在用户设备 (UE, user equipment) 和数据网络 (DN, data network) 之间建立多条协议数据单元 (PDU, protocol data unit) 会话连接，以冗余链路方式保障通信可靠性。但随着 PDU 会话数量增加，用户面流量负载增大，可能造成网络拥塞，而且第三代合作伙伴计划 (3GPP, 3rd generation partnership project) 5G 核心网标准并没有为 PDU 会话制定负载均衡策略或标准。因此，为了提升 5G 核心网服务质量，保证用户面转发效率，本文拟设计并实现一个 5G 核心网用户面流量调度系统。利用 INT 技术实时高精度感知用户面状态信息，并基于该感知信息提出一种改进的路由算法，最终实现 PDU 会话高效转发。

具体而言，本文主要研究内容如下。

1) 提出基于 INT 的 5G 核心网用户面网络状态信息感知方案。创新性地将 INT 技术引入 5G 核心网网络测量中，通过将网络遥测信息插入 GTP-U (GPRS tunnelling protocol user plane) 报文中，实现用户面状态随路测量。同时，基于 INT 技术规范设计了遥测报文格式，保障遥测信息的添加不影响正常数据流量传输。

2) 提出基于改进蚁群算法的流量调度算法。优化传统蚁群算法的信息素初始化方式，提出基于最大最小蚁群系统和精英蚂蚁策略的信息素更新方式。基于感知的网络信息计算启发式函数和信息素函数，实现用户面路由路径规划。

3) 设计实现兼容网络状态信息感知和路由决策的 5G 核心网用户面流量调度系统。利用高精度、细粒度的实时感知信息辅助规划用户面 PDU 会话路由路径。该系统与 Free5GC 的联合部署测试表明，

其不仅能有效感知 5G 核心网用户面相关网络状态信息，而且流量调度算法较传统算法实现了更低的端到端时延和丢包率，并确保负载均衡。

## 1 系统组成

### 1.1 系统定位

应用功能 (AF, application function) 为 5G 核心网提供额外数据，影响 5G 核心网的策略控制，以及边缘计算应用等高级服务<sup>[13]</sup>。根据 AF 是否与 5G 核心网网元处于同一可信域中 (由运营商决定)，AF 与 5G 核心网的交互有 2 种模式，如图 1 所示。

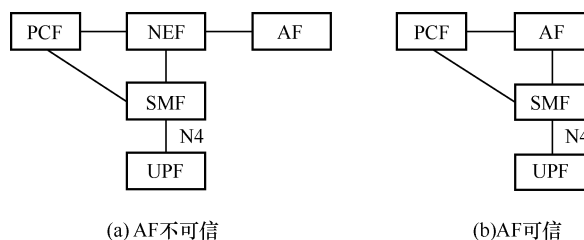


图 1 AF 与 5G 核心网交互模式

当 AF 不可信时 (AF 是由第三方服务提供商提供的服务)，其必须通过网络开放功能 (NEF, network exposure function) 提供的接口将相关信息传输给其他网元，例如接入和移动性管理功能 (AMF)、会话管理功能 (SMF)、策略控制功能 (PCF) 以及统一数据管理功能 (UDM)。当 AF 可信时 (AF 是网络运营商内部开发的功能)，其可以通过服务化接口直接与其他网络通信。所提流量调度系统将作为 5G 核心网控制面中的一个 AF，其主要功能是对 UE 发起的 PDU 会话进行流量调度，同时考虑到当前开源 5G 核心网 (Free5GC、Free5GS 等) 均未实现 NEF 相关功能，所以本文将所提系统作为一个可信 AF，并将其部署于 5G 核心网控制面。

### 1.2 流量调度系统架构

5G 核心网通过建立 PDU 会话提供 UE 与 DN 间的数据连接，针对 PDU 会话规模急剧膨胀导致会话策略相互影响、网络拥塞，以及缺乏负载均衡策略等问题，本文依据“观察、判断、决策、执行”的网络控制环路思想<sup>[14]</sup>，设计面向 5G 核心网用户面的流量调度系统，主要包含用户面网络状态信息感知与路由决策 2 个子系统，如图 2 所示。

网络状态信息感知子系统和路由决策子系统之间通过 RESTful 应用程序接口 (API, application programming interface) 交互，两者模块化的设计降

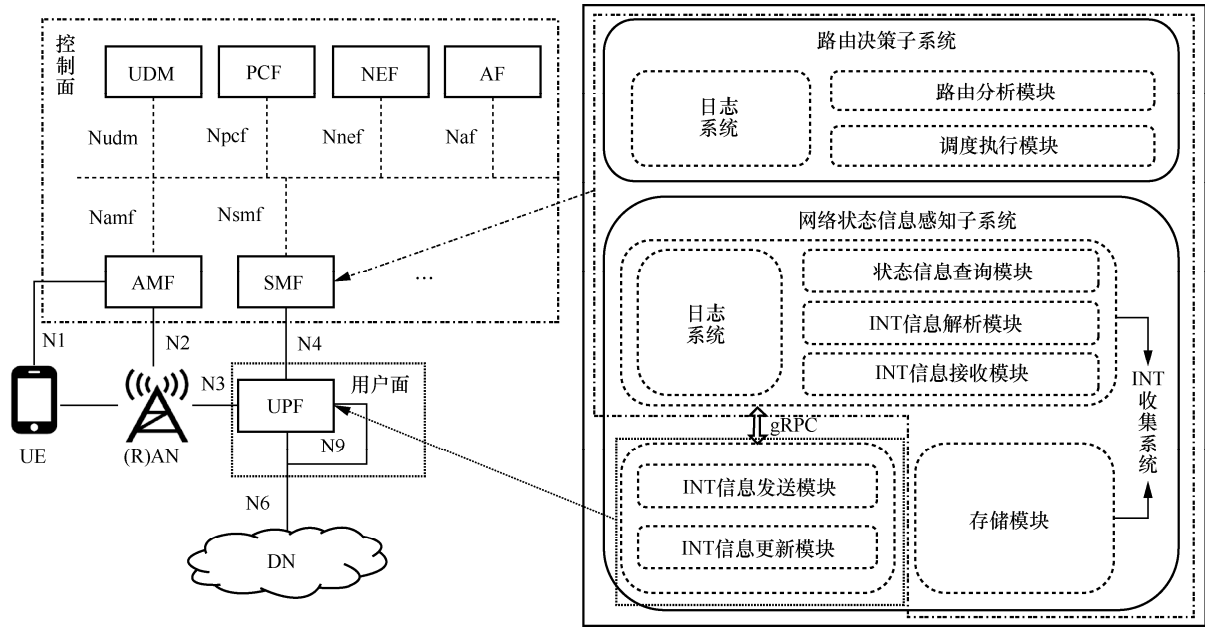


图 2 面向 5G 核心网用户面的流量调度系统总体架构

低了系统耦合性，并增强了系统的可扩展性。所提系统内部模块根据部署位置的不同分为控制面模块和用户面模块。流量调度系统部署于 5G 核心网控制面，直接与 SMF 通信，为其提供流量调度功能。流量调度系统用户面模块部署于 UPF 上，以实现 5G 核心网用户面网络状态信息的实时感知。其中，网络状态信息感知子系统是路由决策子系统运行的基础，它通过 INT 感知并收集用户面网络状态信息，如 UPF 间传输时延、UPF 内处理时延、UPF 端口带宽等。网络状态信息感知子系统内部采用 gRPC<sup>[15]</sup>实现位于控制面的模块与位于用户面模块之间遥测信息的传输。路由决策子系统负责执行流量调度算法，确定当前 PDU 会话最佳路由路径，通过 SMF 更新相应 UPF 的包检测规则(PDR, packet

detection rule)、操作转发规则(FAR, forwarding action rule)，最终将决策下发给 UPF。具体来说，控制面模块包含整个路由决策子系统以及网络状态信息感知子系统部分功能；用户面模块包含 INT 信息发送模块和 INT 信息更新模块。各模块具体功能如表 1 所示。

网络状态信息感知子系统和路由决策子系统分别解决核心网用户面的状态测量和路由调度问题，2 个子系统的功能相互补充、配合，最终实现高精度、细粒度的 5G 核心网用户面网络状态信息感知，并依据状态信息对新建立的 PDU 会话进行路由决策，最终提高核心网服务质量。

### 1.3 流量调度系统流程

面向 5G 核心网用户面的流量调度系统运行流程如图 3 所示。网络状态信息感知子系统在系统启

表 1

模块功能

模块	功能
INT 信息更新模块	根据报文采样频率、INT 指令、UPF 角色等参数，对接收到的 GTP-U 报文进行 INT 报头插入、INT 元数据更新、INT 信息提取等
INT 信息发送模块	将收集到的 INT 信息发送给数据收集器并发送相应控制指令给 INT 信息更新模块
INT 信息接收模块	使用 gRPC 接收 INT 信息发送模块传输的 INT 信息，将其解析为 INT 原始报文
INT 信息解析模块	根据预先定义的 INT 报文格式，将 INT 原始报文各字段解析为对应含义的网络状态信息并转换成相应的数据结构
状态信息查询模块	根据其他系统或模块的调用请求，向存储模块查询对应网络状态信息并返回查询结果
存储模块	存储 INT 原始报文和解析后的网络状态信息，实现对网络状态信息的持久化存储
路由分析模块	接收用户面网络状态信息，运行相应的流量调度算法，实现 PDU 会话路由路径规划
调度执行模块	将路由决策发送给 SMF，由 SMF 根据路由决策更新 UPF 的 PDR、FAR，并执行决策
日志系统	保存系统运行日志，方便运维人员了解系统运行状态

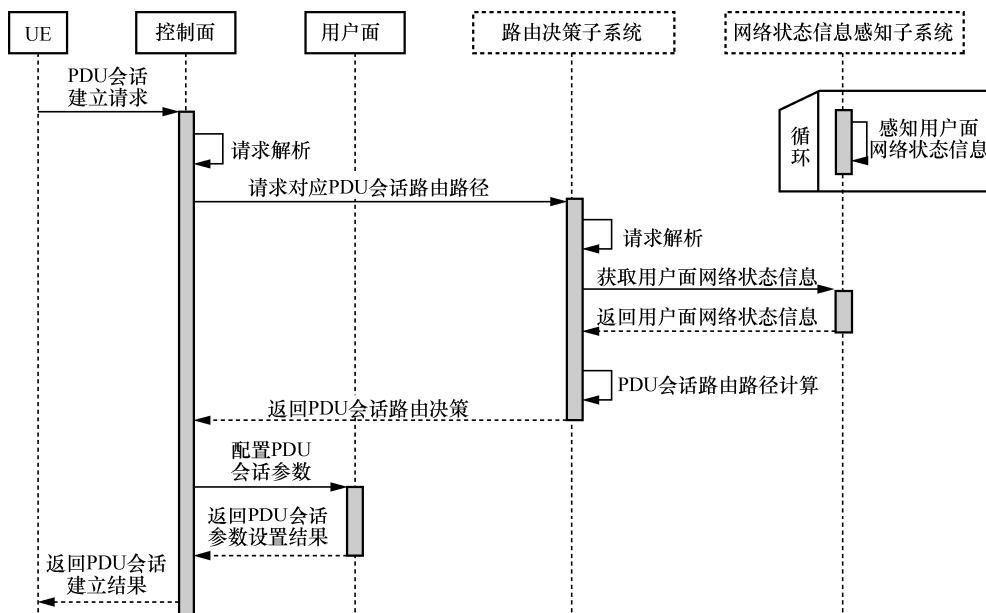


图 3 面向 5G 核心网用户面的流量调度系统运行流程

动时便持续实时感知用户面网络状态信息。

UE 向 5G 核心网控制面发起 PDU 会话建立请求。控制面解析 PDU 会话建立请求后确定 PDU 会话相关参数，如 DN 名称、UE 位置信息等，并向路由决策子系统请求 PDU 会话路由路径。路由决策子系统解析到达的请求，确定 PDU 会话的源节点和目的节点，之后向网络状态信息感知子系统请求当前的用户面网络状态信息。路由决策子系统基于状态信息执行流量调度算法，并将 PDU 会话路由决策返回给控制面。控制面解析该路由决策，生成相应的 PDR、FAR，并使用包转发控制协议(PFCP, packet forwarding control protocol)向 UPF 发送调度指令。用户面将调度结果反馈给控制面，最终由控制面将 PDU 会话建立结果返回给 UE。

#### 1.4 存储机制设计

INT 元数据经解析后的网络状态信息，标识每个 INT 原始报文的唯一标识符等相关数据均需保存在数据库中。INT 原始报文是由作为 INT 接收节点的 UPF 传输给 INT 收集系统的 INT 信息，包括 INT 报文头部和所历经的各个 UPF 插入的 INT 元数据，其大小受到当前流量带宽、测量频率、测量参数等因素影响。本文系统采用 MySQL 数据库实现可靠的数据存储和高效的数据写入。此外，当用户查询当前网络状态信息时，流量调度系统需快速响应以保证数据的实时性，因此，本文系统引入轻量级内存数据库 Redis 作为系统热点数据缓存，使用

Key-Value 的方式存储用户最近查询的热点数据，实现对用户请求的快速响应。为防范因 SQL 注入攻击导致网络状态信息数据库泄露的风险，利用权限设置防止高风险语句执行，同时严格检查输入变量的类型和格式，防止 SQL 注入。针对外部用户恶意访问接口导致的系统能力下降的问题，可采用 Token 授权认证或者黑白名单等方式实现鉴权。

## 2 基于带内网络遥测的状态信息感知

### 2.1 网络状态信息感知子系统

网络状态信息感知子系统作为路由决策子系统的数据库提供方，通过 INT 技术实现对 5G 核心网用户面状态信息的感知，并直接与 UPF 进行交互。由于该子系统需要与 Free5GC 联合部署，而 Free5GC 中 UPF 的 GTP-U 报文处理模式是作为 Linux 内核模块开发使用的，因此为使 UPF 支持 INT 功能，位于用户面的 INT 信息更新模块作为 Linux 内核模块进行开发，INT 信息发送模块位于 Linux 用户空间，采用 Linux 系统中通用的 Netlink 功能<sup>[16]</sup>实现 2 个模块之间的通信。

基于 INT 的网络状态信息感知流程如图 4 所示。其中，INT 源节点是指创建 INT 报头 (INT Header) 并将其插入报文的 UPF，INT 中间节点是指根据 INT 报头中包含的指令收集用户面遥测信息 (元数据) 的 UPF，INT 接收节点是指提取 INT 报头并收集遥测信息的锚点 UPF (PSA-UPF, PDU

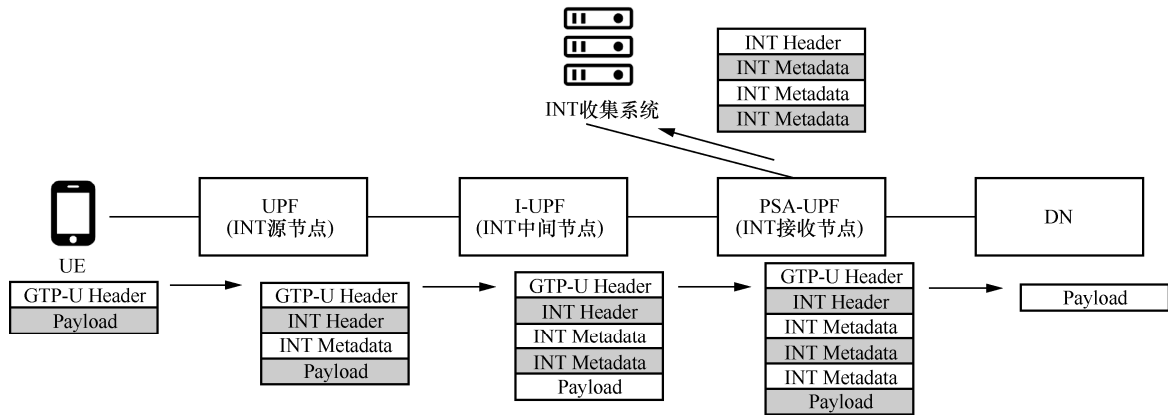


图 4 基于 INT 的网络状态信息感知流程

session anchor-UPF), INT 头部是 INT 遥测信息的报头, INT 元数据 (INT Metadata) 是需要收集的网络状态信息。

### 2.2 INT 报文结构

5G 核心网用户面中承载用户报文的主要协议为 GTP-U, 该协议为支持未来的应用程序, 实现了一些可用于传输任意类型数据的 GTP-U 扩展头部, 例如在 3GPP 5G 核心网标准中已定义了几个扩展值: PDU 会话容器、服务等级指示器等。完整 GTP-U 报文格式请参考文献[17]。

本文将所有 INT 信息, 包括 INT 报文、INT 指令、INT 元数据, 均放入 GTP-U 扩展头部, 并将其作为一个新的 GTP-U 扩展报文类型。若 GTP-U 经过的部分路由器、交换机、UPF 不支持 INT, 该部分网络设备或 UPF 能够自动忽略 GTP-U 扩展报头中的 INT 信息并自动转发报文, 保证数据流量的正常转发以及网络性能的准确测量。本文参考 P4.org 应用工作组提出的 INT 规范<sup>[18]</sup>, 结合 5G 核心网用户面状态信息感知需求, 设计图 5 所示的 INT 报文格式。

协议版本号 (4 bit) 固定为 2; D (1 bit) 代表丢弃标志位, 该值为 1 时, INT 接收节点在读取报文 INT 元数据后不转发该报文并丢弃; E (1 bit)

代表超过最大跳数标志位, 当 INT 节点不能添加元数据且剩余跳数为 0 时, 该值为 1; M (1 bit) 代表超过最大传输单元 (MTU, maximum transmission unit) 限制标志位; Reserved (12 bit) 代表保留位, 在 INT 源节点该值为 1, 在其余 INT 节点该字段被忽略; Hop ML (5 bit) 代表 INT 元数据长度; Remaining Hop Count (8 bit) 代表允许添加 INT 元数据的剩余跳数; Instruction Bitmap (16 bit) 是 INT 指令位图, 每个比特位表示一种网络遥测数据类型; Domain Specific ID (16 bit) 代表 INT 域 ID, 规定当前 INT 指令适用的范围, INT 域 ID 为 0 时, 表示所有 INT 节点均能根据 INT 指令插入对应 INT 元数据。

INT 报头共 8 B, 之后 INT 报文内容为 INT 元数据协议栈。INT 元数据的插入顺序受到 INT 指令位图的严格限制, 以方便 INT 收集系统根据数据类型解析对应的 INT 元数据。该 INT 数据插入方式实现了较强的兼容性, 当 UPF 不支持 INT 时, 网络转发设备可以自动忽略 GTP-U 扩展头部中包含的不能识别的 INT 信息, 继续进行 GTP-U 报文转发, 逐步替换原有不支持 INT 的 UPF 设备, 最终实现基于 INT 技术的 5G 核心网用户面网络状态信息感知。



图 5 INT 报文格式

### 2.3 INT 报文处理

INT 报文在 INT 源节点、INT 中间节点和 INT 接收节点中的处理方式如下。

#### 1) INT 源节点

INT 源节点根据预先设定的 INT 指令和测量频率对接收到的 GTP-U 报文进行封装，包括 INT 报头和 INT 元数据，生成并转发带有 INT 信息的 GTP-U 报文。作为 INT 源节点的 UPF 收到 GTP-U 报文后，首先判断 GTP-U 报文是否满足预先设定的测量频率（设定测量频率为  $2^f$ ，其中  $f$  为 8 bit 的无符号十进制整数），若不满足则视为普通 GTP-U 报文进行转发；若满足则进入添加 INT 信息的流程。INT 源节点确认需向 GTP-U 报文添加 INT 信息后，首先检查该 GTP-U 报文是否有足够空间容纳 INT 信息，若空间充足，则向 GTP-U 扩展头部中添加 INT 报头，并将 INT 源节点元数据插入 INT 元数据栈中；若没有足够空间容纳 INT 信息或报文长度超过 MTU，则单独构造 GTP-U 格式的遥测报文，同时向该报文添加 INT 信息。之后将丢弃指示位 D 置 1，表示该报文为包含 INT 信息的遥测报文，需在 INT 接收节点丢弃。最后，更新 GTP-U 报文上层协议校验和（如 IP 报头校验和），通过 UPF 的 PDR、FAR 转发至下一跳 UPF。

#### 2) INT 中间节点

INT 中间节点负责采集本地 UPF 的 INT 元数据，并将其插入具有 INT 报头的 GTP-U 报文的 INT 元数据栈中。作为 INT 中间节点的 UPF 首先对收到 GTP-U 报文进行解析，若该报文不是包含 INT 信息的遥测报文，则根据 PDR、FAR 进行转发；若该报文是带 INT 信息的遥测报文，则根据 INT 指令判断 GTP-U 报文是否有足够空间或当前长度是否超过 MTU。如果空间不足或长度超过 MTU，则同样单独构建 GTP-U 格式的遥测报文，并将原 GTP-U 报文中的 INT 报头和所有 INT 元数据转移到新构建的 GTP-U 报文中，同时将丢弃指示位 D 置 1；如果空间充足且长度不超过 MTU，则将本地 UPF 的 INT 元数据插入 INT 元数据栈中。最后，更新 GTP-U 报文上层协议校验和（如 IP 报头校验和），按照 UPF 的 PDR、FAR 进行转发。

#### 3) INT 接收节点

INT 接收节点除向 GTP-U 报文添加 INT 信息外，主要负责提取 GTP-U 报文中的 INT 报头和 INT 元数据，并通过 gRPC 将 INT 信息发送给 INT 收集系统。

作为 INT 接收节点的 UPF 收到 GTP-U 报文后进行解析，若该报文非遥测报文，则按照 PDR、FAR 进行转发；若该报文为遥测报文，可按照 INT 信息添加流程向该 GTP-U 报文添加 INT 信息，之后提取所有 INT 信息，将其通过 gRPC 发送给 INT 收集系统。若该 GTP-U 报文非 INT 探测包，即丢弃标志位 D 不为 1，则删除包含的所有 INT 信息，同时更新 GTP-U 上层协议校验和，根据 PDR、FAR 进行转发；否则丢弃该 INT 探测包。

### 2.4 基于 Netlink 的数据传输方案设计

在网络状态信息感知子系统实现层面，5G UPF 的 GTP-U 报文解析与转发功能是以 Linux 内核模块的方式实现的，对应程序运行在 Linux 内核空间。UPF 转发用户面 GTP-U 报文时，所有 GTP-U 报文均在 Linux 内核空间中实现解析与转发。因此，网络状态信息感知子系统向 GTP-U 报文中添加 INT 报头和 INT 元数据的功能也必须在 Linux 内核空间实现，但 gRPC 框架只有处于 Linux 用户空间时才能使用。所以，本文系统基于 Netlink 通信框架将位于 Linux 内核空间的 INT 信息更新模块采集到的 INT 信息传输给位于 Linux 用户空间的 INT 信息发送模块。

网络状态信息感知子系统实现的适用于 INT 的 Netlink 子协议号为 31。该数据传输方案中，从 Linux 用户空间向 Linux 内核空间发送的消息为控制指令，主要功能是设置 UPF 所对应的 INT 角色，设置测量频率和需要感知的网络状态信息类型，从 Linux 内核空间向 Linux 用户空间发送的消息为 UPF 作为 INT 接收节点提取到的所有 INT 信息。

位于用户空间的 INT 信息发送模块和位于内核空间的 INT 信息更新模块之间的消息交互过程如图 6 所示。

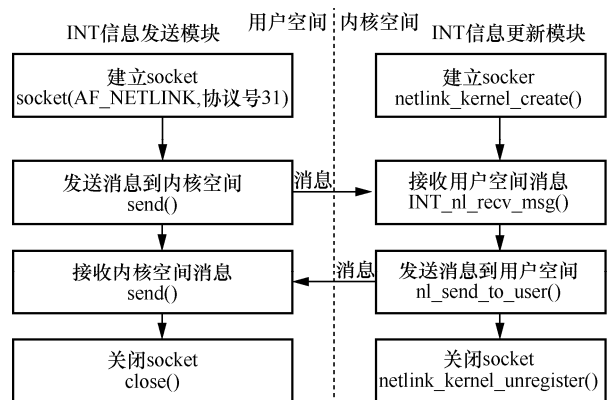


图 6 消息交互过程

在该数据传输方案中,首先加载 INT 信息更新模块,使其创建内核空间的 Netlink 套接字,INT 信息更新模块在内核空间采用 netlink\_unicast()函数对指定进程标识符(PID, process identifier)的用户空间进程进行通信,所以需要 INT 信息发送模块率先使用 send()函数发送初始消息到内核空间,使 INT 信息更新模块能够获取 PID。

### 3 基于改进蚁群算法的路由决策

#### 3.1 路由决策子系统

路由决策子系统是 5G 核心网用户面流量调度系统的控制中枢,其与网络状态信息感知子系统交互获取 5G 核心网用户面状态信息,如 UPF 端口带宽、UPF 间链路时延、UPF 内处理时延等,通过运行流量调度算法,生成 PDU 会话路由决策,保障不同用户服务的 QoS。同时路由决策子系统直接与 5G 核心网控制面的 SMF 交互,获取流量调度请求并将流量调度决策下发。路由决策子系统逻辑架构如图 7 所示。

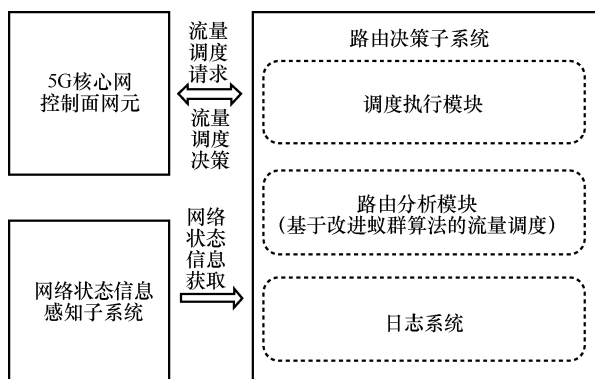


图 7 路由决策子系统逻辑架构

#### 3.2 路由决策子系统算法设计

路由决策子系统的目的是通过运行流量调度和路由算法,在 5G 核心网用户面现有时延和带宽等条件约束下,为新建的 PDU 会话规划一条合理路由路径,避免网络拥塞的同时提高用户面通信链路利用率与 PDU 会话吞吐量,降低 PDU 会话时延。本文借鉴蚁群算法在寻找最优路径方面的分布式计算、鲁棒性强、不需要全局信息等特点<sup>[19-20]</sup>,将 5G 核心网用户面当前网络状态信息(链路时延、UPF 吞吐量等)加入蚂蚁寻路过程中。在已有的基于蚁群算法的网络路由相关工作中,网络状态信息是固定配置的,蚁群算法不能实时动态感知变化的网络环境。本文提出一种基于改进蚁群算法的流量调度(IAC-TS, traffic scheduling based on improved

ant colony)算法,利用 INT 状态信息感知系统为蚁群算法提供动态网络信息,在一次迭代中每一只蚂蚁单独对从基站到 DN 的路径进行搜索,形成备选路径集合。之后,更新蚂蚁经过路径的信息素并开始下一次迭代。随着迭代次数增加,蚂蚁会集中到信息素浓度较高的几条路径上,直至到达指定迭代次数,获取最优路由路径集合。实现过程如下。

##### 1) 信息素初始化

使用常数对信息素初始化的方式容易导致收敛速度过慢,为避免该问题,将网络链路的信息素初始值设为

$$\tau_{ij}(0) = B_{ij}^{\max} - \frac{T x_{ij}}{T} \quad (1)$$

其中,  $\tau_{ij}(0)$  表示初始链路  $(i, j)$  上的信息素含量,在 5G 核心网用户面流量调度背景下,该值表示链路的可用带宽;  $B_{ij}^{\max}$  表示链路  $(i, j)$  的最大可用带宽;  $T x_{ij}$  表示链路  $(i, j)$  在时间  $T$  内发送的字节数。

##### 2) 蚂蚁转移

蚂蚁的初始位置均位于作为源节点的 UPF 上。规定蚂蚁路径不能形成环路,因此每只蚂蚁都会维护一个禁忌表,该表存放蚂蚁经过 UPF 节点 ID。蚂蚁依据链路信息素浓度计算转移到下一个节点的概率,计算方式为

$$P_{ij}^k(t) = \begin{cases} \frac{|\tau_{ij}(t)|^\alpha [\eta_{ij}(t)]^\beta}{\sum_{u \in N_k^i} [\tau_{iu}(t)]^\alpha [\eta_{iu}(t)]^\beta}, & j \in N_k^i \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中,  $i$  和  $j$  分别表示蚂蚁的当前节点和下一跳节点,  $\tau_{ij}(t)$  表示时刻  $t$  链路  $(i, j)$  上的信息素浓度,  $\eta_{ij}(t)$  表示时刻  $t$  链路  $(i, j)$  上的启发信息,  $\alpha$  和  $\beta$  分别表示信息素浓度和启发信息的权重因子,  $N_k^i$  表示位于节点  $i$  的蚂蚁  $k$  所有未经过节点的集合。

$\eta_{ij}(t)$  计算式为

$$\eta_{ij}(t) = \frac{1}{\sigma_{ij}(t)} \quad (3)$$

$$\sigma_{ij}(t) = \gamma \frac{D_{ij}(t) - D_{\min}(t)}{D_{\max}(t) - D_{\min}(t)} + (1 - \gamma) \frac{B_{ij}(t)}{B_{ij}^{\max}} \quad (4)$$

其中,  $\sigma_{ij}(t)$  是链路  $(i, j)$  的质量权衡因子,  $D_{ij}(t)$  和  $\frac{B_{ij}(t)}{B_{ij}^{\max}}$  分别是链路  $(i, j)$  的时延和带宽利用率,  $D_{\min}(t)$  和  $D_{\max}(t)$  分别是所有链路中的时延最小值

和最大值,  $\gamma(0 < \gamma < 1)$  是权重因子。

若蚂蚁所有相邻节点均在禁忌表中, 则该蚂蚁的此次路径搜索失败。否则, 蚂蚁通过轮盘赌方式选择下一跳节点, 同时检查禁忌表中是否包含源节点和目的节点, 若包含则搜索成功。否则, 蚂蚁重新计算转移概率并选择下一跳节点。

### 3) 信息素更新

所有蚂蚁执行完一次路径搜索, 即完成一次迭代后, 将所得有效路径通过 Top- $k$  算法选择最优的  $k$  条路径添加到候选路径集中, 之后, 根据式(5)更新当前链路信息素浓度

$$\tau_{ij}(t + \Delta t) = (1 - \rho)\tau_{ij}(t) \quad (5)$$

其中,  $\tau_{ij}(t + \Delta t)$  是经过时间  $\Delta t$  后链路  $(i, j)$  的信息素浓度,  $\rho(0 < \rho < 1)$  是信息素挥发因子。为提升蚁群算法前期的搜索能力和后期的收敛速度,  $\rho$  的取值方式为

$$\rho = \begin{cases} 0.9, & 0 \leq I < 0.25I_{\max} \\ 0.5, & 0.25I_{\max} \leq I < 0.75I_{\max} \\ 0.1, & 0.75I_{\max} \leq I < I_{\max} \end{cases} \quad (6)$$

其中,  $I$  表示算法的当前迭代次数,  $I_{\max}$  表示最大迭代次数。

之后再次更新候选路径集中  $k$  条链路的信息素浓度以增加优质链路的信息素浓度, 保证链路质量, 即

$$\tau_{ij}(t + \Delta t) = \tau_{ij}(t) + (1 + \mu_{ij})\Delta\tau_{ij} \quad (7)$$

$$\Delta\tau_{ij} = \sum_{n=1}^m \Delta\tau_{ij}^n(t) \quad (8)$$

$$\Delta\tau_{ij}^n(t) = \begin{cases} \frac{Q}{L_n}, & \text{第 } n \text{ 只蚂蚁经过链路 } (i, j) \\ 0, & \text{其他} \end{cases} \quad (9)$$

$$\mu_{ij} = \begin{cases} \mu, & e_{ij} \in P \\ 0, & \text{其他} \end{cases} \quad (10)$$

其中,  $\Delta\tau_{ij}$  表示链路  $(i, j)$  新增的信息素浓度;  $\Delta\tau_{ij}^n(t)$  表示第  $n$  只蚂蚁在链路  $(i, j)$  上释放的信息素浓度;  $m$  表示蚁群大小;  $L_n$  表示第  $n$  只蚂蚁走过的路径总长度, 以路径总时延表示;  $Q$  表示信息素总浓度, 为一常量;  $e_{ij}$  表示以  $i, j$  为端点的 UPF 链路;  $P$  表示候选路径集;  $\mu$  表示候选路径的信息素增量因子。

为防止链路信息素浓度出现极端情况导致蚂蚁搜索路径不准确, 基于最大最小蚁群思想对信息素增量  $\Delta\tau_{ij}$  进行限制, 即

$$\tau_{ij}(t + \Delta t) = \begin{cases} \tau_{\max}, & \tau_{ij}(t + \Delta t) > \tau_{\max} \\ \tau_{ij}(t + \Delta t), & \tau_{\min} \leq \tau_{ij}(t + \Delta t) \leq \tau_{\max} \\ \tau_{\min}, & \tau_{ij}(t + \Delta t) < \tau_{\min} \end{cases} \quad (11)$$

其中,  $\tau_{\max}$ 、 $\tau_{\min}$  分别表示链路  $(i, j)$  的最大、最小信息素浓度。为模拟现实网络状态, 本文仿真实验将可用带宽限制在  $[5, 12]$  Mbit/s。

### 3.3 基于改进蚁群算法的流量调度方法

算法执行到最大迭代次数  $I_{\max}$  后, 在候选路径集中随机选择一条链路作为 PDU 会话的最终路由路径, 算法的具体实现过程如算法 1 所示。

#### 算法 1 IAC-TS 算法

输入  $I_{\max}$ ,  $\alpha$ ,  $\beta$ ,  $\mu$ , 蚂蚁数量 num, 蚂蚁索引 ID, 网络拓扑  $G$ , 链路时延, 带宽, 源节点 src, 目标节点 dest

输出 PDU 路由路径

1) Func ANT\_COLONY\_OPT( $G$ , src, dest)

2) 创建可选路径集合

3) for  $i = 1, 2, \dots, I_{\max}$

4) for  $j = 1, 2, \dots, \text{num}$

5) 利用蚁群算法寻找通路

6) if src, dest  $\in$  path

7) 将所得通路添加到可选路径集合

8) end if

9) end for

10) 获得最优的  $k$  条通路

11) for  $j = 1, 2, \dots, \text{num}$

12) 更新信息素浓度

13) end for

14) end for

15) 获得最优路径

16) return chosen\_path

17)end Func

18)Func FIND\_PATH(ID,  $G$ , src, dest)

19) 创建禁忌表

20) 向禁忌表添加源节点

21) for curr\_loc  $\neq$  dest

22) 更新下一跳节点链表

23) 计算并选择概率最大的下一跳节点

24) 将所选节点添加到路径集合

25) 蚂蚁位置转移

26) 向禁忌表添加当前节点

27) end for

### 28)end Func

算法 1 中 1)~17)行用于计算获取最优路径, 每只蚂蚁在成功搜索到从源节点到目的节点的一条链路后, 对所有链路信息素浓度进行更新, 采用 Top-k 算法选出 k 条候选路径。18)~28)行表示单只蚂蚁搜索最优路径的过程, 采用轮盘赌算法选择下一跳节点, 并用式(2)计算下一跳节点的选择概率。

## 4 系统验证

### 4.1 实验环境

本文所提系统旨在实现 5G 核心网用户面流量优化调度, 具体地, 通过 INT 技术对用户面网络状态信息感知, 并基于路由决策子系统的路由算法实现 PDU 会话路由路径规划。将所提 5G 核心网用户面流量调度系统与 Free5GC 进行联合部署, 测试环境如表 2 所示。

表 2 测试环境

设备	规格
处理器	AMD Ryzen R5 3 600
内存	64 GB
硬盘	1 TB
显卡	GeForce RTX 2 060 Super
功能	系统名称
操作系统	Ubuntu 20.04
虚拟化平台	Docker v20.10.22
5G 核心网	Free5GC v3.0.6
UE、RAN 模拟器	UERANSIM v3.2.6
流量生成器	Iperf v2.0
网络流量控制	TC
数据库	MySQL 8.0
缓存系统	Redis v6.2

### 4.2 网络状态信息感知功能验证

网络状态信息感知子系统负责监控 5G 核心网用户面网络状态, 通过 INT 技术实时获取网络状态信息, 包括用户面链路时延、UPF 内 GTP-U 报文处理时延、UPF 端口吞吐量等。网络状态信息感知子系统功能测试拓扑如图 8 所示。

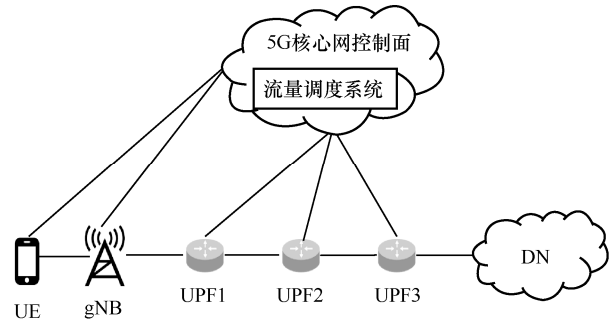


图 8 网络状态信息感知子系统功能测试拓扑

该网络拓扑包括一个 5G 基站 (gNB)、3 个具有 INT 功能的 UPF、一个基于 Docker 部署的 DN, 以及 5G 核心网控制面 (包含完整的核心网功能、路由决策子系统功能、部分网络状态信息感知子系统功能)。在 UE 中运行 Iperf 以模拟 PDU 会话建立请求, UPF1-UPF3 分别作为 INT 源节点、INT 中间节点、INT 接收节点, 最终将报文发送至 DN。

#### 4.2.1 INT 功能测试

利用 UERANSIM 建立一条从 UE 到 DN 的 PDU 会话路径, 在 UE 上发送测试报文并使用 Wireshark 抓包, 以验证 GTP-U 报文中是否携带 INT 信息, 测试结果如图 9 所示。

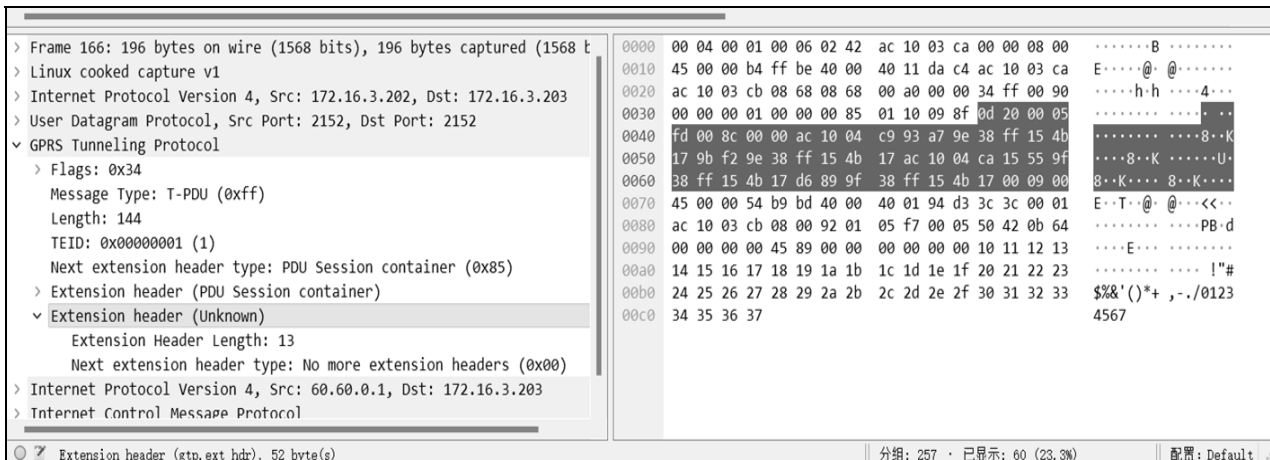


图 9 测试结果

测试结果显示，UPF1 向该报文的 GTP-U 扩展头部中插入了 INT 报头和 INT 元数据，UPF2 作为中间节点仅插入了 INT 元数据，UPF3 将本地设备收集的 INT 元数据和 GTP-U 报文中的所有 INT 信息一同发给位于 5G 控制面的 INT 收集系统。该测试结果说明 INT 功能可以正常使用。

#### 4.2.2 不同遥测参数对吞吐量的影响

本节实验测试不同遥测参数对 PDU 会话吞吐量产生的影响。实验选用 3 种典型的 INT 参数的不同组合进行测试：RX+TX 表示测量报文经过 UPF 端口的接收/发送的字节数，Egress TS+Ingress TS (Timestamp) 表示测量报文经过 UPF 出口/入口的时间戳，None 表示未启用 INT 功能。不同遥测参数组合对吞吐量的影响如图 10 所示，其中遥测间隔是指向每  $n$  个报文中插入一项 INT 元数据。

未启用 INT 功能时，PDU 会话带宽是一个定值。随着遥测间隔减小，即测量频率升高，不同遥测参数组合的 PDU 会话带宽都会明显降低，而且更多遥测参数的组合会在测量频率升高时使 PDU 会话带宽快速下降，原因是过高的测量频率会导致在报文中插入了大量的 INT 信息，UPF 内部处理时延明显增加，UPF 端口拥塞，从而导致吞吐量下降。因此，过多的遥测参数会对 PDU 会话性能产生不利影响。

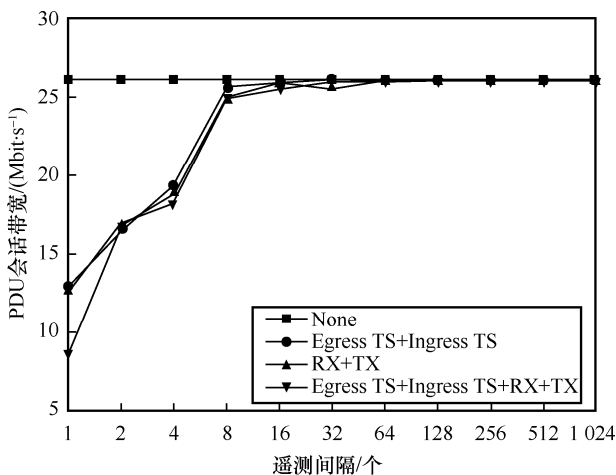


图 10 不同遥测参数组合对吞吐量的影响

#### 4.2.3 不同遥测参数对时延的影响

通过分析用户面 GTP-U 报文中的 INT 元数据，获取 UPF1 到 UPF3 的单向传输时延。不同遥测参数组合对传输时延的影响如图 11 所示。

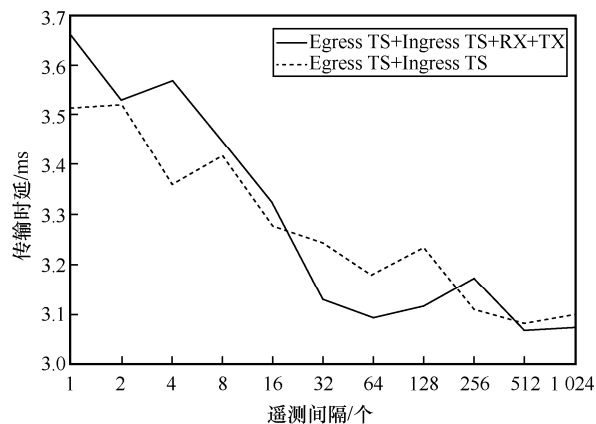


图 11 不同遥测参数组合对传输时延的影响

从图 11 可以看出，在不同的遥测参数组合下，测量频率会影响 PDU 会话时延。遥测间隔越小，即测量频率越高，传输时延越大，这也与 GTP-U 中的 INT 信息处理有关。因此在实际部署 INT 功能时，需要对测量精度与对 PDU 会话质量产生的影响进行权衡。

#### 4.2.4 不同遥测参数对 UPF CPU 利用率的影响

本节实验测试不同遥测参数在不同遥测间隔下对 UPF3，即 INT 接收节点的 CPU 利用率的影响，在 UPF3 上监控该设备的 CPU 利用率，实验结果如图 12 所示。

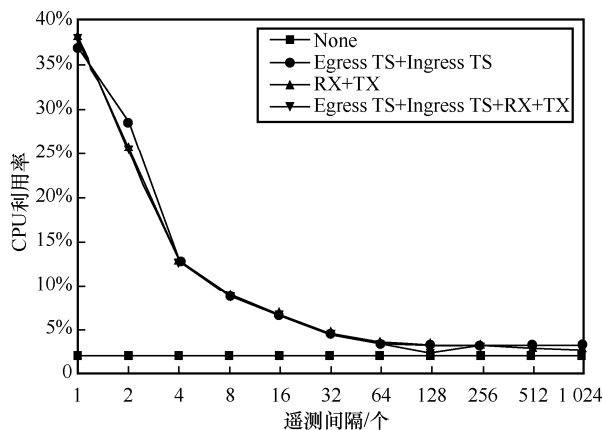


图 12 不同遥测参数在不同遥测间隔下对接收节点 CPU 利用率的影响

随着遥测间隔减小，即测量频率提升，不同遥测参数组合的 UPF3 的 CPU 利用率都快速提升。一方面，INT 接收节点设备本身在收集 INT 信息和提取 INT 信息时会占用一定 CPU；另一方面，系统采用的 Netlink 和 gRPC 技术频繁向 INT 收集系统发送 INT 信息也会占用 CPU。因此，需

要选择合理的测量频率,将 INT 引入对 CPU 的额外占用的影响降到最小。

### 4.2.5 链路负载变化对 UPF 处理时延的影响

本节实验通过改变 PDU 路径上各个 UPF 的背景流量大小,测量 UPF 单跳处理时延来测试 INT 的效果。将链路带宽设置为 30 Mbit/s,背景流量设置为 40 Mbit/s,目的是模拟链路中存在拥塞的情况。在 UE 上通过 Iperf 向 DN 发送 UDP 报文,测量不同遥测间隔对 UPF3 单跳时延的影响,实验结果如图 13 和图 14 所示。

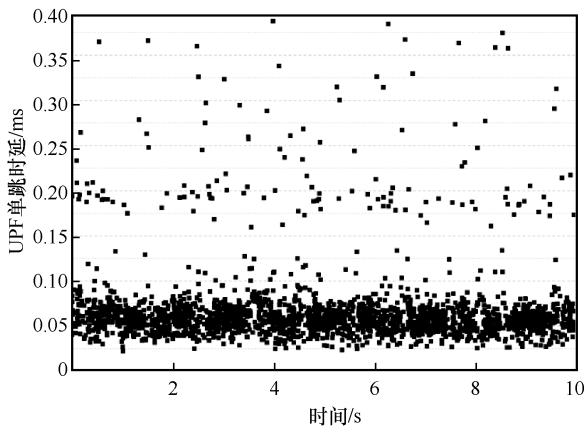


图 13 UPF 单跳时延 (遥测间隔为一个报文)

在遥测间隔减小,即遥测频率提升的情况下,所探测的 GTP-U 报文数量增加,因此 UPF 的处理开销增大,使 GTP-U 报文转发时延增加。另外,由于设置的背景流量大于链路带宽,因此会有部分 GTP-U 报文的处理时延偏高。

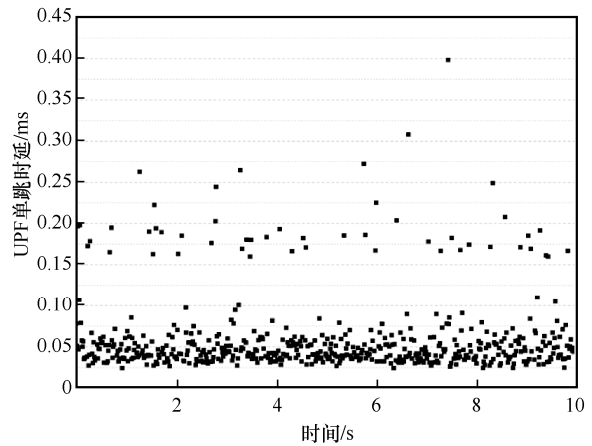


图 14 UPF 单跳时延 (遥测间隔为 8 个报文)

### 4.3 流量调度功能验证

本节对由网络状态信息感知子系统和路由决策子系统组成的整个流量调度系统进行测试。首先路由决策子系统接收 SMF 发送的 PDU 会话路由请求;然后从网络状态信息感知子系统感知用户面网络状态信息,计算 PDU 会话路由路径;最后将路由决策返回给 SMF,由 SMF 下发路由决策。流量调度系统功能测试拓扑如图 15 所示。用户面共包含 11 个 UPF,UPF1 作为 INT 源节点,UPF9-UPF11 作为 INT 接收节点,其余 UPF 均为 INT 中间节点。

#### 4.3.1 PDU 会话流量调度功能测试

本节实验测试路由决策子系统与 SMF 之间的连通性。将所提流量调度系统与 Free5GC 联合部署,通过 UERANSIM 模拟的 UE 发起 PDU 会

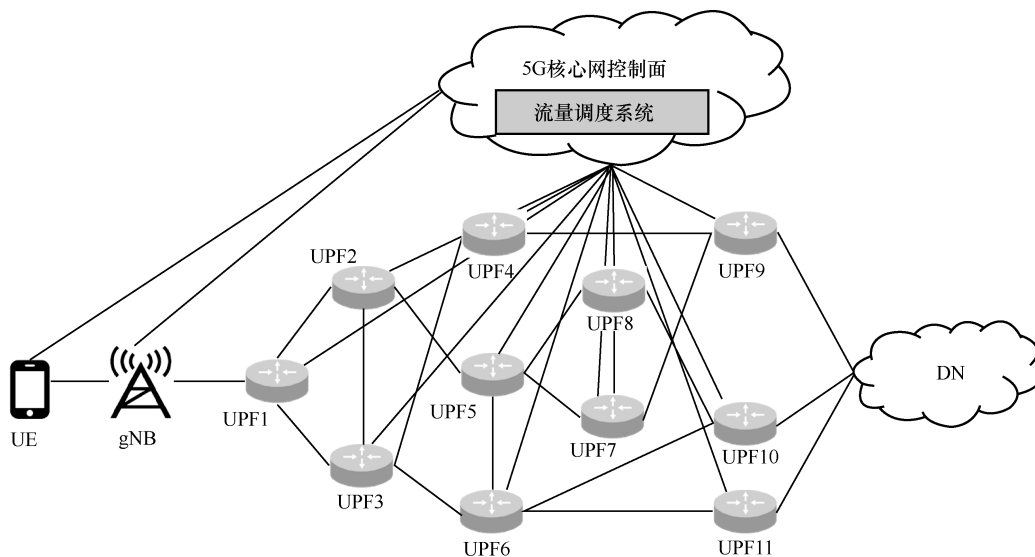


图 15 流量调度系统功能测试拓扑

话建立请求，发送数据报文进行测试，在网络状态信息数据库中查询报文中的 INT 元数据，通过跟踪该报文的路径信息，判断是否执行正确的路由决策。

图 16 显示了运行流量调度系统后，UE 建立 PDU 会话连接时的 5G 核心网运行结果。

由图 16 可知，在 UE 请求建立 PDU 会话过程中，SMF 向流量调度系统请求对应 PDU 会话的路

由路径，最终完成 PDU 会话建立。解析完成后的 INT 元数据如图 17 所示。

图 17 显示了 2 个携带 INT 信息的 GTP-U 报文经过不同的 UPF，data\_id 为 3801616~3801618 是第一个报文，data\_id 为 3801619~3801621 是第二个报文，node\_id 是 GTP-U 经过的不同 UPF 的编号，结果显示路由决策子系统选用了不同的 UPF 对 GTP-U 报文进行了转发。

```

2023-03-09T13:58:47Z [DEBU][SMF][PduSess] Install SessionRule[SessRuleId-1]
2023-03-09T13:58:47Z [INFO][SMF][PduSess] SUPI[imsi-20893000000003] use LoadBalancer to determine the route path
2023-03-09T13:58:47Z [INFO][SMF][LoadBalancer] LB selected path id[2], path[172.16.4.201172.16.4.209172.16.4.204172.16.4.209172.16.4.212]
2023-03-09T13:58:47Z [INFO][SMF][Consumer] SendNFDiscoveryServingAMF ok
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionEstablishmentResponse
2023-03-09T13:58:47Z [INFO][LIB][PCF] Remove Request Transaction [13]
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionEstablishmentResponse
&{200 Mbps 100 Mbps}
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionEstablishmentResponse
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionEstablishmentResponse
2023-03-09T13:58:47Z [INFO][LIB][PCF] Remove Request Transaction [15]
2023-03-09T13:58:47Z [INFO][LIB][PCF] Remove Request Transaction [17]
2023-03-09T13:58:47Z [INFO][LIB][PCF] Remove Request Transaction [14]
2023-03-09T13:58:47Z [INFO][LIB][PCF] Remove Request Transaction [16]
2023-03-09T13:58:47Z [INFO][SMF][GIN] | 201 | 172.16.4.102 | POST | /nsmf-pdusession/v1/sm-contexts |
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionEstablishmentResponse
2023-03-09T13:58:47Z [INFO][SMF][PduSess] Recieve Update SM Context Request
2023-03-09T13:58:47Z [INFO][SMF][PduSess] In HandlePDUSessionSMContextUpdate
2023-03-09T13:58:47Z [INFO][SMF][PCF] In HandlePfcPSessionModificationResponse
2023-03-09T13:58:47Z [INFO][SMF][PduSess] [SMF] PCF Modification Resonse Accept
2023-03-09T13:58:47Z [INFO][SMF][PCF] PCF Session Modification Success[1]
    
```

图 16 5G 核心网运行结果

data_id	node_id	egress_time	ingress_time	timestamp	flow_id	hop_latency	packet_id	queue_id	queue_occupancy	rx	tx
3801616	-922153044	1668093709538173447	1668093709538144547	2022-11-10 23:21:50	0	0	985675784527616	0	0	593118700	578766344
3801617	-888598612	1668093709550810964	1668093709550794667	2022-11-10 23:21:50	0	0	985675784527616	0	0	296547724	296565288
3801618	-771158100	1668093709575959805	1668093709575929805	2022-11-10 23:21:50	0	0	985675784527616	0	0	75322	90342
3801619	-922153044	1668094222332312561	1668094222332297261	2022-11-10 23:30:23	0	0	985680087883520	0	0	593193564	578857756
3801620	-905375828	1668094222345412358	1668094222345394358	2022-11-10 23:30:23	0	0	985680087883520	0	0	288352622	288316076
3801621	-771158100	1668094222370744117	1668094222370725717	2022-11-10 23:30:23	0	0	985680087883520	0	0	151208	181504

图 17 解析完成后的 INT 元数据

### 4.3.2 不同算法对 PDU 会话流量调度性能的影响

PDU 会话建立完成后，在 DN 中设置 Iperf 服务端程序测量 PDU 会话端到端时延、带宽以及丢包率，UE 通过 Iperf 客户端发送 UDP 报文模拟数据流量。设置 UE 每隔 10 s 发起一个新的 PDU 会话请求，每个成功建立的 PDU 会话均发送带宽固定为 1.05 Mbit/s 的 UDP 报文，发送报文持续时长为 200 s，直至建立 10 条 PDU 会话。将所提 IAC-TS 算法与传统路由算法（包括最短路径算法、ECMP 算法，以及广度优先搜索（BFS, breadth-first search）算法）进行性能对比，实验结果如图 18~图 20 所示。

从图 18~图 20 可以看出，所提 IAC-TS 算法在网络负载相同的情况下，PDU 会话的端到端传输时延和丢包率均低于最短路径算法、ECMP 算法和 BFS 算法，而平均吞吐量具有明显提升。具体来看，较其他算法，IAC-TS 算法的端到端传输时延平均降低 72.03%，端到端平均吞吐量平均提升 9.88%，

端到端平均丢包率降低 76.1%。说明基于改进蚁群算法的路由方案使路由决策子系统能够做出更好的 PDU 会话路由决策，避免用户面网络拥塞，提升 PDU 会话吞吐量和带宽，达到了负载均衡效果，进而提升了服务质量。

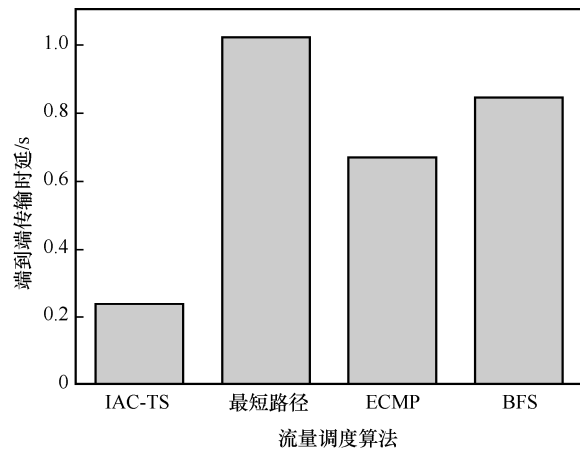


图 18 端到端传输时延

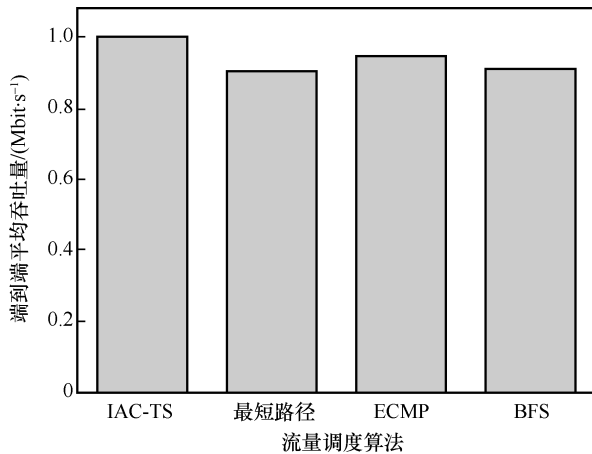


图 19 端到端平均吞吐量

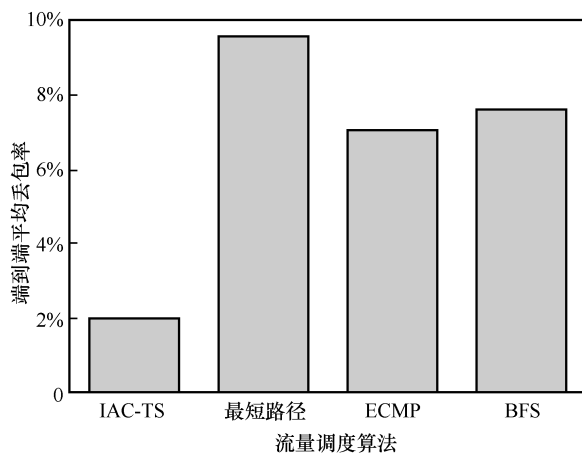


图 20 端到端平均丢包率

## 5 结束语

5G 复杂应用和庞大业务流量对核心网服务质量提出更高需求。本文设计并实现了 5G 核心网用户面流量调度系统，包含网络状态信息感知子系统和路由决策子系统。网络状态信息感知子系统采用 INT 实现高精度、细粒度的用户面状态信息感知，路由决策子系统依据用户面状态信息，执行基于改进蚁群算法的流量调度算法，生成 PDU 会话路由路径。在测试阶段，将所提流量调度系统与 Free5GC 联合部署，证明了所提系统能够以较小的网络开销实现对用户面状态信息的感知，而且所做 PDU 会话路由决策在时延、吞吐量和丢包率等方面均优于其他几种路由决策算法。

在未来工作中，将进一步改进 INT 信息收集架构，避免因测量频率提升导致 INT 信息占用大量带宽的问题；此外，需要实现测量频率、网络状态参数自动调整以节约测量开销并快速定位网络问题；

还将考虑采用更加先进高效的流量调度算法来提升 5G 核心网用户面数据流量的路由效果。

## 参考文献:

- [1] 张平, 牛凯, 田辉, 等. 6G 移动通信技术展望[J]. 通信学报, 2019, 40(1): 141-148.  
ZHANG P, NIU K, TIAN H, et al. Technology prospect of 6G mobile communications[J]. Journal on Communications, 2019, 40(1): 141-148.
- [2] ITU. IMT traffic estimates for the years 2020 to 2030[R]. 2015.
- [3] TAKAHASHI K, AIDA K, TANJO T, et al. A portable load balancer with ECMP redundancy for container clusters[J]. IEICE Transactions on Information and Systems, 2019, E102.D(5): 974-987.
- [4] MA S Q, TANG H, WANG X X. A link load balancing algorithm based on ant colony optimization in data center network[C]//Proceedings of the International Conference on Communications and Networking in China. Berlin: Springer, 2020: 474-489.
- [5] SILVA L S D, STORCK C R, DUARTE-FIGUEIREDO F D L P. A dynamic load balancing algorithm for data plane traffic[C]//Proceedings of the Latin American Network Operations and Management Symposium. Piscataway: IEEE Press, 2019: 1-7.
- [6] TAN T J, WENG F L, HU W T, et al. A reliable intelligent routing mechanism in 5G core networks[C]//Proceedings of the 26th Annual International Conference on Mobile Computing and Networking. New York: ACM Press, 2020: 1-3.
- [7] NGUYEN V G, GRINNEMO K J, TAHERI J, et al. On auto-scaling and load balancing for user-plane gateways in a softwarized 5G network[C]//Proceedings of the International Conference on Network and Service Management. Piscataway: IEEE Press, 2021: 132-138.
- [8] LAI S Q, YUAN X L, LIU J K, et al. OblivSketch: oblivious network measurement as a cloud service[C]//Proceedings the Network and Distributed System Security Symposium. Virginia: the Internet Society, 2021: 1-18.
- [9] 吕鸿润, 李清, 沈耿彪, 等. 带内网络遥测方法综述[J]. 软件学报, 2023, 34(8): 3870-3890.  
LYU H R, LI Q, SHEN G B, et al. Survey on In-band network telemetry[J]. Journal of Software, 2023, 34(8): 3870-3890.
- [10] YAO H P, MAI T L, XU X B, et al. NetworkAI: an intelligent network architecture for self-learning control strategies in software defined networks[J]. IEEE Internet of Things Journal, 2018, 5(6): 4319-4327.
- [11] ISOLANI P H, HAXHIBEQIRI J, MOERMAN I, et al. An SDN-based framework for slice orchestration using in-band network telemetry in IEEE 802.11[C]//Proceedings of the 6th IEEE Conference on Network Softwarization (NetSoft). Piscataway: IEEE Press, 2020: 344-346.
- [12] KATTA N, GHAG A, HIRA M, et al. Clove: congestion-aware load balancing at the virtual edge[C]//Proceedings of the 13th International Conference on emerging Networking Experiments and Technologies. New York: ACM Press, 2017: 323-335.
- [13] LENTISCO C M, BELLIDO L, CÁRDENAS A, et al. Design of a 5G

multimedia broadcast application function supporting adaptive error recovery[J]. IEEE Transactions on Multimedia, 2023, 25: 378-388.

- [14] 祝学军, 赵长见, 梁卓, 等. OODA 智能赋能技术发展思考[J]. 航空学报, 2021, 42(4): 10-19.
- ZHU X J, ZHAO C J, LIANG Z, et al. Thoughts on technology development of OODA empowered with AI[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 10-19.
- [15] BISWAS R, LU X Y, PANDA D K. Accelerating TensorFlow with adaptive RDMA-based gRPC[C]//Proceedings of IEEE 25th International Conference on High Performance Computing (HiPC). Piscataway: IEEE Press, 2019: 2-11.
- [16] SALIM J H, KHOSRAVI H, KLEEN A, et al. Netlink as an IP services protocol[J]. RFC, 2003: doi.org/10.17487/RFC3549.
- [17] 3GPP. General packet radio system (GPRS) tunnelling protocol user plane: 29.281[S]. 2020.
- [18] P4.org. In-band network telemetry (int) dataplane specification[Z]. (2020-11-11)
- [19] RATHEE M, KUMAR S, GANDOMI A H, et al. Ant colony optimization based quality of service aware energy balancing secure routing algorithm for wireless sensor networks[J]. IEEE Transactions on Engineering Management, 2021, 68(1): 170-182.
- [20] XIANG X S, TIAN Y, ZHANG X Y, et al. A pairwise proximity learning-based ant colony algorithm for dynamic vehicle routing problems[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(6): 5275-5286.

#### [作者简介]



王超 (1996- ), 男, 山东潍坊人, 北京邮电大学博士生, 主要研究方向为 5G/6G 核心网、卫星网络。



任蒙 (1997- ), 男, 重庆人, 北京邮电大学硕士生, 主要研究方向为 5G/6G 核心网。



王尚广 (1982- ), 男, 河南周口人, 博士, 北京邮电大学教授、博士生导师, 主要研究方向为服务计算、边缘计算与智能、卫星计算、5G/6G 核心网等。