

文章编号: 2095-4980(2024)09-0983-09

基于缓存的卫星传输性能优化策略

陆涣冰^a, 任壮壮^a, 丁晓进^b, 张更新^b

(南京邮电大学 a.物联网学院; b.通信与信息工程学院, 江苏 南京 210003)

摘要: 传统卫星透明转发模式在新型星地网络下会导致巨大的时延。为提高对地面用户的服务能力, 本文在低轨卫星和地面内容服务器上设置缓存, 并联合低轨卫星、地面内容服务器以及地面用户在同频干扰下, 针对星地网络的传输吞吐问题进行建模。使用改进的粒子群(PSO)算法和支持超平面优化(SHOT)求解器求解。实验表明, 对目标优化问题求解时, SHOT 求解器获得的曲线较为平滑, 优化的 PSO 算法相比 SHOT 求解器得到的结果, 有 2% 上下的波动。在信干噪比阈值为 8 dB 和最大功率 20 W, 以及卫星或地面内容服务器干扰数量分别为 2 和 3 的约束下, 可分别求得 45 Mbps 和 37 Mbps 的系统吞吐量。在分别设置低轨卫星、地面内容服务器以及地面用户数量的情况下, 本文对两种算法的运行时间进行了分析, 最坏情况下设置 3 类数量为 4 时, 优化 PSO 解法比 SHOT 求解器快约 10 s。结果表明, 启发式算法更适用于未来复杂星地网络建模, 更适合寻找非凸非线性全局解, 并在高维问题下求解速度表现更优。

关键词: 低轨卫星; 星地网络; 缓存; 粒子群; SHOT 求解器

中图分类号: TN927+.2

文献标志码: A

doi: 10.11805/TKYDA2023056

Performance optimization strategies for satellite transmission based on caching

LU Huanbing^a, REN Zhuangzhuang^a, DING Xiaojin^b, ZHANG Gengxin^b

(a.School of Internet of Things; b.School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210003)

Abstract: The traditional satellite transparent relay mode causes significant delays in the new satellite-ground network. In order to improve the service capability for ground users, this paper proposes a caching scheme on both the Low-Earth Orbit(LEO) satellite and the ground content server, and models the transmission throughput problem of the satellite-ground network considering co-frequency interference among LEO satellites, ground content servers, and ground users. The improved Particle Swarm Optimization(PSO) algorithm and the Supporting Hyperplane Optimization Toolkit (SHOT) solver are employed to solve the problem. The experimental results reveal that the SHOT solver can obtain a smoother result curve, and there is only about 2% fluctuation compared to the optimization PSO algorithm and the SHOT solver. Under the constraint of an 8 dB signal-to-noise ratio threshold and a maximum power of 20 W, the proposed approach can achieve a transmission rate of 45 Mbps and 37 Mbps, respectively, with interference from 2 satellites or 3 ground content servers. Furthermore, the study analyzes the running time of the two algorithms for various numbers of low-earth orbit satellites, ground content servers, and ground users. The optimization PSO algorithm is approximately 10 s faster than the SHOT solver, particularly when the number of the three categories is set to 4, in the worst-case scenario. The results indicate that heuristic algorithms are more suitable for modeling future complex satellite-ground networks and finding non-convex and nonlinear global solutions. They also perform better in terms of solving speed for high-dimensional problems.

Keywords: Low-Earth Orbit satellites; satellite-ground network; cache; Particle Swarm Optimization; Supporting Hyperplane Optimization Toolkit

低轨卫星和卫星组网技术是目前卫星和航天领域的研究热点, 这种技术可以提供高速、广覆盖的通信服务,

收稿日期: 2023-03-04; 修回日期: 2023-04-27

满足现代人类对信息的需求^[1]。近年来,随着卫星技术的不断发展和应用需求的增长,卫星组网得到了广泛的研究和关注。如,研究者通过将多个卫星连接成网络,形成卫星组网,提高了卫星的可靠性和覆盖率^[2];也有研究者将卫星组网技术用于地球观测,提高地球观测的精确度和效率^[3]。

传统低轨卫星或卫星组网本身使用人造卫星作为中继站,在地面内容站和卫星站之间进行信号的高速转发工作。但随着5G、卫星物联网、卫星互联网以及云计算技术和卫星硬件的飞速发展,传统低轨卫星只作为中继和转发的能力逐渐显现短板,星地之间频繁的冗余交互,导致了性能的巨大损失;暴露出星上数据链路阻塞、星地组网系统出现瓶颈、星间传输能力薄弱等诸多问题,对下一代的低轨卫星在全球通信体系中的能力提出了更高的要求,需在传统低轨卫星高速中继转发的能力上,提高就地存储、提供热点缓存以及就地计算等多样能力。

本文在上述背景下,在低轨卫星上设置缓存,并联合地面内容服务器的缓存,进行协同优化,提高低轨卫星-地面内容服务器-地面用户的整体吞吐,旨在为地面用户提供响应更快的服务。在上述模型下,本文对整个系统进行建模和分析。根据星地和地面两个通信场景的不同信道特性,对两个阶段进行建模:从低轨卫星到地面内容服务器的信道模型为莱斯信道;从地面内容服务器直接到地面用户的信道为瑞利信道。通过设置指示变量对两个阶段不同的信道模型进行统一,结合布置的缓存的内容流行度,对时间周期内的星地通信吞吐做出优化;使用优化的启发式算法(PSO算法)和SHOT求解器分别求解星地间的缓存优化问题,均能达到最优解;并针对系统的性能部分单独优化,对缓存、文件到达和排队时间做出更细致的分析。

1 系统模型和问题建模

如图1所示,系统场景由低轨卫星部分 LEO_j 和地面内容服务器 CS_i 以及地面用户 GU_k 三部分组成,其中 J 表示场景中低轨卫星总数, I 表示地面内容服务器总数, K 表示地面用户总数。具体的,第 j 颗卫星表示为 LEO_j , $1 \leq j \leq J$;第 i 个地面内容服务器表示为 CS_i , $1 \leq i \leq I$;第 k 个地面用户表示为 GU_k , $1 \leq k \leq K$ 。简单起见,设 LEO_j 到 CS_i 的下行链路为LC链路; CS_i 到 GU_k 的下行链路为CG链路。

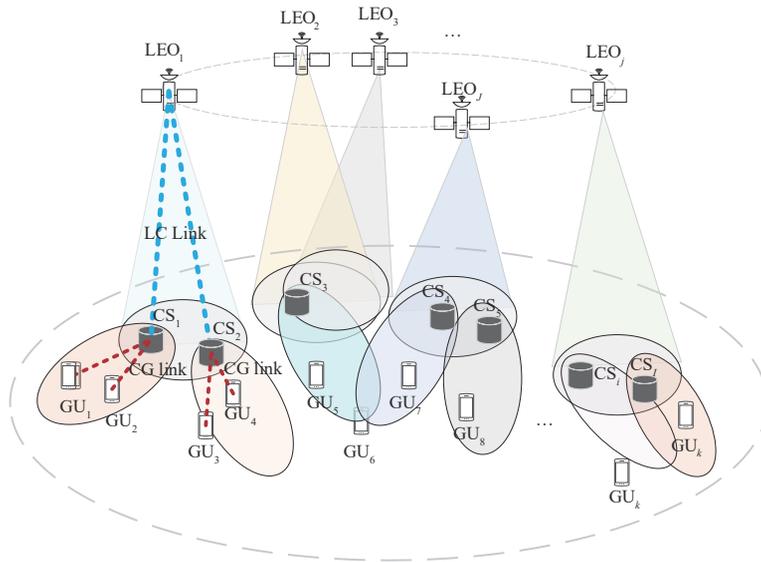


Fig.1 Scene map
图1 场景图

在时隙 t 内, GU_k 向 CS_i 发送文件 f 的请求表示为 $GU_{k,f}^{CS_i}$, CS_i 收到请求后,在本地的缓存中检查是否命中文件缓存。如果命中,则直接下发给 GU_k ;如果未命中文件 f ,则 CS_i 向 LEO_j 请求文件。 CS_i 收到 LEO_j 的回复后,计算文件流行度,判断是否缓存在当前本地缓存中。为提高吞吐,并减少回程链路上的拥堵问题,地面内容服务器使用了缓存,可以在非高峰时段预先缓存一段来自低轨卫星的服务内容,然后将这些信息转发给地面用户。为节省预算,没有在地面用户侧部署高增益天线,此外,考虑到地面内容服务器的基站集群会因地理距离或同频发生干扰,同理相同的低轨卫星星座群中也会因为频率的相同或不同发生干扰,因此也针对星间频率干扰做出考量。

1.1 用户内容请求模型

由于低轨卫星飞行时间有限,为便于分析,将时间分为 T 个长度相等的时间间隔 δt 。由于 δt 选得足够小,因

此可认为地面内容服务器和地面用户的距离在每个时间间隔内不变。

设内容总体为 F , $F = \{1, 2, 3, \dots, f, \dots, F\}$, 所有的内容假设为 Q bit 大小, 如果有不同大小的内容, 可切分为相同大小的块。本文采用 Zipf 的流行度分布模型, 可得到地面用户 GU_k 访问任意内容 f 的流行度模型:

$$P_{GU_k f} = \frac{f^{-\vartheta}}{\sum_{i=1}^F i^{-\vartheta}} \quad (1)$$

式中: $f^{-\vartheta}$ 表示本次文件被访问, $\vartheta \in (0, 1)$ 为 Zipf 分布的偏度因子; $\sum_{i=1}^F i^{-\vartheta}$ 表示除 f 之外被访问的文件访问总和; $\sum P_{GU_k f} = 1$ 。

本文假设地面内容服务器可以在业务非高峰时段主动将部分热门内容缓存在其存储结构中。值得注意的是, 缓存放置需在地面内容服务器上线前离线执行。不失一般性, 定义 $\eta = \{\eta_f, \forall f \in F\}$ 为地面内容服务器的二元缓存放置变量, 其中 $\eta_f \in \{0, 1\}$ 。具体地, $\eta_f = 1$, 表明地面内容服务器缓存了热点内容 f , 地面用户 GU_k 可直接从地面内容服务器的缓存中获取 f ; $\eta_f = 0$, 则表明地面内容服务器中并没有缓存内容 f , 需要地面内容服务器向 LEO 卫星请求内容, 以服务地面用户。一旦地面内容服务器收到 GU_k 的请求, 将立即开始向用户传输内容, 且地面内容服务器配备了 MQ bit 的缓存空间, 其中 M 为地面内容服务器可存储的最大内容数量。由于缓存容量有限, 地面内容服务器无法存储所有的文件内容, 因此有 $M < F$, 每个 GU 都会独立地约束内容。考虑上述有限的缓存容量, 给出缓存约束:

$$\sum_{f=1}^F \eta_f \leq M \quad (2)$$

1.2 LC 链路通信模型

设 $x_{LC,LEO}^t$ 为在单位功率下, 时隙 t 中, 从目标 LEO 传输到地面的符号, 并有 $E[|x_{LC,LEO}^t|^2] = 1$ 。考虑地面内容服务器时隙 t 中和目标低轨卫星通信时, 存在其他的同频星地间通信干扰。因此, 考虑目标星地之间的通信模型为:

$$y_{LC,CS_i}^t = \sqrt{p_{LC,LEO_j}^t g_{LC}^t(d_{LEO_j,CS_i}) G_{LC,LEO_j,CS_i}^t G_{LC,LEO_j}^t h_{LC,LEO_j}^t x_{LC,LEO_j}^t} + \sum_{x=1, x \neq j}^{\phi} \sqrt{p_{LC,LEO_x}^t g_{LC}^t(d_{LEO_x,CS_i}) G_{LC,LEO_x,CS_i}^t G_{LC,LEO_x}^t h_{LC,LEO_x}^t x_{LC,LEO_x}^t} + n_{LEO,CS_i} \quad (3)$$

式中: ϕ 表示干扰信号集合; p_{LC,LEO_j}^t 表示在时隙 t 中, 在 LC 链路上, 目标 LEO_{*j*} 到 CS_{*i*} 的传输功率; p_{LC,LEO_x}^t 表示干扰信号中在时隙 t 下 LC 链路上目标 LEO_{*x*} 到 CS_{*i*} 的传输功率; $g_{LC}^t(d_{LEO_j,CS_i})$ 表示在时隙 t 内, LC 链路上 LEO_{*j*} 到 CS_{*i*} 的自由空间大尺度衰落; $g_{LC}^t(d_{LEO_x,CS_i})$ 表示干扰信号中, 在时隙 t 内 LC 链路 LEO_{*x*} 到 CS_{*i*} 的自由空间大尺度衰落; G_{LC,LEO_j,CS_i}^t 、 G_{LC,LEO_j}^t 分别表示时隙 t 下, LC 链路中 CS_{*i*} 的接收增益和 LEO_{*j*} 的发射增益; G_{LC,LEO_x,CS_i}^t 、 G_{LC,LEO_x}^t 分别表示时隙 t 下, LC 链路干扰信号中 CS_{*i*} 的接收增益和 LEO_{*x*} 的发射增益; h_{LC,LEO_j}^t 和 h_{LC,LEO_x}^t 分别为 LC 链路下, 符合莱斯信道的小尺度衰落; n_{LEO,CS_i} 表示 CS_{*i*} 接收到均值为 0, 方差为 $kB_{LC}T_n$ 的加性高斯白噪声, 其中 k 、 B_{LC} 、 T_n 分别表示玻尔兹曼常数、LC 链路上的带宽以及等效噪声温度。

可得到 LC 链路上信干噪比(Signal to Interference plus Noise Ratio, SINR)为:

$$R_{SIN,LC}^t = \frac{p_{LC,LEO_j}^t g_{LC}^t(d_{LEO_j,CS_i}) G_{LC,LEO_j,CS_i}^t G_{LC,LEO_j}^t |h_{LC,LEO_j}^t|^2}{\sum_{x=1, x \neq j}^{\phi} p_{LC,LEO_x}^t g_{LC}^t(d_{LEO_x,CS_i}) G_{LC,LEO_x,CS_i}^t G_{LC,LEO_x}^t |h_{LC,LEO_x}^t|^2 + kB_{LC}T_n} \quad (4)$$

最后可得到 t 时隙下 LC 链路上 CS_{*i*} 接收 LEO_{*j*} 文件回复的速率:

$$r_{LC}^t = B_{LC} \log_2(1 + R_{SIN,LC}^t) \quad (5)$$

1.3 CG 链路通信模型

设 $x_{CG,CS}^t$ 为在单位功率下, 时隙 t 中, 从目标 CS 到地面用户传输的符号, 并有 $E[|x_{CG,CS}^t|^2] = 1$ 。考虑到时隙 t 中, 地面用户和地面内容服务器通信时, 存在其他同频通信干扰, 因此通信模型为:

$$y'_{CG, GU_k} = \sqrt{p'_{CG, CS_i} g'_{CG}(d_{CS_i, GU_k}) G'_{CG, CS_i, GU_k} G'_{CG, CS_i} h'_{CG, CS_i}} \cdot x'_{CG, CS_i} + \sum_{v=1, v \neq i}^{\varphi} \sqrt{p'_{CG, CS_v} g'_{CG}(d_{CS_v, GU_k}) G'_{CG, CS_v, GU_k} G'_{CG, CS_v} h'_{CG, CS_v}} \cdot x'_{CG, CS_v} + n \quad (6)$$

式中： p'_{CG, CS_i} 表示在时隙 t 中，在CG链路上，目标 CS_i 到 GU_k 的传输功率； p'_{CG, CS_v} 表示干扰信号中，时隙 t 下CG链路上目标 CS_v 到 GU_k 的传输功率； $g'_{CG}(d_{CS_i, GU_k})$ 表示在时隙 t 内，CG链路上 CS_i 到 GU_k 的自由空间大尺度衰落； $g'_{CG}(d_{CS_v, GU_k})$ 表示干扰信号中，时隙 t 内CG链路 CS_v 到 GU_k 的自由空间大尺度衰落； G'_{CG, CS_i, GU_k} 、 G'_{CG, CS_i} 分别表示时隙 t 下，CG链路中 GU_k 的接收增益和 CS_i 的发射增益； G'_{CG, CS_v, GU_k} 、 G'_{CG, CS_v} 分别表示时隙 t 下，CG链路干扰信号中 GU_k 的接收增益和 CS_v 的发射增益； h'_{CG, CS_i} 和 h'_{CG, CS_v} 为CG链路下符合瑞利信道的小尺度衰落，且 $E[|h'_{CG, CS_i}|^2] = 1$ ； n 为 CS_i 接收到均值为0，方差为 σ^2 的高斯白噪声。

将瑞利信道小尺度衰落 h'_{CG, CS_i} ^[4]进行展开：

$$h'_{CG, CS_i} = \sqrt{\frac{k}{1+k}} \bar{h}'_{CG, CS_i} + \sqrt{\frac{1}{1+k}} \hat{h}'_{CG, CS_i} \quad (7)$$

式中： \bar{h}'_{CG, CS_i} 和 \hat{h}'_{CG, CS_i} 分别为视距(Line of Sight, LOS)和非视距(Non Line of Sight, NLOS)的信道成分； k 为瑞利因子，是固定路径功率和散射路径功率的比值，当 $k=0$ 时，表示没有固定路径功率的接收路径。一般来说，当地面内容服务器和地面用户没有遮挡时，信道模型为LOS，此时衰减少。和NLOS相比，LOS信道模型的信号质量更好，吞吐量更大。而当两端之间有建筑、植物等遮挡时，信道模型选用NLOS，并且此时还有建筑物等遮挡物的信号反射、衍射和穿透损耗。

综上，可得到CG路径上的信干噪比为：

$$R'_{SIN, CG} = \frac{p'_{CG, CS_i} g'_{CG}(d_{CS_i, GU_k}) G'_{CG, CS_i, GU_k} G'_{CG, CS_i} |h'_{CG, CS_i}|^2}{\sum_{v=1, v \neq i}^{\varphi} p'_{CG, CS_v} g'_{CG}(d_{CS_v, GU_k}) G'_{CG, CS_v, GU_k} G'_{CG, CS_v} \cdot |h'_{CG, CS_v}|^2 + \sigma^2} \quad (8)$$

最后可得 t 时隙下CG链路上 GU_k 接收 CS_i 文件回复的速率表达式为：

$$r'_{CG} = B_{CG} \log_2(1 + R'_{SIN, CG}) \quad (9)$$

式中 B_{CG} 为 CS_i 和 GU_k 之间的带宽。

1.4 优化问题建模

综上所述，可得地面用户 GU_k 在时间 T 内可以获得的总吞吐 Y ：

$$Y = \delta_t P_{GU_k, f} \left[\eta_f \sum_{t_{LC}}^T r'_{LC} + (1 - \eta_f) \sum_{t_{CG}}^T r'_{CG} \right] \quad (10)$$

式中： δ_t 为单位时间； $P_{GU_k, f}$ 为地面内容服务器中的缓存命中率； η_f 和 $1 - \eta_f$ 分别表示某一时隙下是否命中地面内容服务器缓存； t_{LC} 为总时间 T 内，卫星缓存文件直接向地面内容服务器发送文件的第一个时隙； t_{CG} 为地面内容服务器向地面用户发送文件内容的第一个时隙； r'_{LC} 和 r'_{CG} 分别表示2个不同阶段的可以达到的最大速率。

目标优化函数和其中各项约束为：

$$\left\{ \begin{array}{l} \max Y \\ p'_{LC, LEO_j}, p'_{CG, CS_i, \eta_f} \\ \left\{ \begin{array}{l} 0 \leq p'_{LC, LEO_j} \leq p_{LEO_{max}} \\ 0 \leq p'_{CG, CS_i} \leq p_{CS_{max}} \\ R'_{SIN, LC} \geq R_{SIN, LC_{min}} \\ R'_{SIN, CG} \geq R_{SIN, CG_{min}} \\ \eta_f \in \{0, 1\}, \forall f \in F \\ \sum_{f \in F} \eta_f \leq M \end{array} \right. \\ \text{s.t.} \end{array} \right. \quad (11)$$

式中： $p_{LEO_{max}}$ 、 $p_{CS_{max}}$ 分别为低轨卫星和地面内容服务器可以达到的最大功率； $R_{SIN, LC_{min}}$ 、 $R_{SIN, CG_{min}}$ 分别为LC链路和

CG链路上任意LEO_j和CS_i的信干噪比阈值。上述优化问题是一个混合整数非线性规划问题，要获得直接解是困难的，为解决这一问题，本文提供了两种解决方案。

2 问题求解

本文提出的目标优化问题是混合非整数非线性规划问题。传统PSO算法无法有效地找到混合整数非线性规划(Mixed-Integer Nonlinear Programming, MINLP)问题的全局最优解，因此本文改进了PSO算法的实现，旨在相同的优化问题约束下，和SHOT求解器给出的结果相互验证，证明解的正确性。

2.1 优化粒子群算法求解

朴素PSO算法^[5]维护粒子集合 $(x_1, x_2, x_3, \dots, x_p)$ ，这些粒子均匀地分布在问题搜索空间的周围。具体来说，设 q 为集合的大小，对于任意粒子 x_i ，算法使用以下方式更新粒子位置：

$$x_{k+1}^i = x_k^i + v_{k+1}^i \quad (12)$$

式中下标 k 表示一轮迭代的时间刻度。

在粒子集中的每个粒子都与一个速度 v_{k+1}^i ($-v_{\max} \leq v_{k+1}^i \leq v_{\max}$)相关联，表示解空间中粒子的位置变化率。

$$v_{k+1}^i = w_k v_k^i + c_1 r_1 (q_k^i - x_k^i) + c_2 r_2 (q_k^g - x_k^i) \quad (13)$$

式中： r_1 和 r_2 为0和1之间的均匀随机数^[6]； c_1 和 c_2 为朴素PSO中粒子的认知参数和社会参数； w_k 为惯性权重，用来约束上一轮的历史速度对当前轮速度的影响，平衡粒子的全局和局部空间搜索的范围。式(13)用来计算粒子的最新速度 v_{k+1}^i ，根据其上一代的速度 v_k^i 和其当前的位置 x_k^i 与其本身局部历史最佳位置 p_k^i 和全局粒子最佳位置 p_k^g 进行更新。

初始化时 $k=0$ ，粒子速度的初始值 v_i^0 为约束 $0 \leq v_0 \leq v_{\max}$ 内的随机整数。具体地， v_{\max} 约束了朴素PSO中粒子的最大全局搜索范围，在不断迭代求出每一轮的速度后，粒子最终可以根据位置变化公式移动到新的位置，尝试找出最优解。

朴素PSO对于解决优化问题是有效的，但其基本运算不适用于处理复杂约束条件，因此本文在朴素PSO的基础上做出修改，在算法流程中修改等式约束，使用惩罚函数法缩小解空间，提升朴素PSO求解问题的能力^[7]。

2.1.1 算法改进

将优化问题抽象表示为：

$$\begin{cases} \max f(\mathbf{x}, \mathbf{y}) \\ \text{s.t.} \begin{cases} [g(\mathbf{x}, \mathbf{y}) \geq 0] \\ [h(\mathbf{x}, \mathbf{y}) = 0] \\ [\mathbf{x} \in \mathbf{X}, \mathbf{y} \in \mathbf{Y}] \end{cases} \end{cases} \quad (14)$$

设向量 \mathbf{X} 和向量 \mathbf{Y} 分别由 n 个连续实数变量 \mathbf{x} 和 m 个整数变量 \mathbf{y} 组成^[8]，有 $\mathbf{X} \subset \mathbf{R}^n$ 和 $\mathbf{Y} \subset \mathbf{N}^m$ 。 $f(\mathbf{x}, \mathbf{y})$ 为目标优化函数， $g(\mathbf{x}, \mathbf{y}) \geq 0$ 和 $h(\mathbf{x}, \mathbf{y}) = 0$ 分别表示解维度为 p 和 q 的不等式约束和等式约束。分解连续变量 \mathbf{x} 和整数 \mathbf{y} 为：

$$\mathbf{x} = [v, \zeta]^T, \mathbf{y} = [w, \xi]^T \quad (15)$$

假设 v 和 w 为等式约束 $h(\mathbf{x}, \mathbf{y}) = 0$ 的独立变量^[9]， $v \in \Gamma \subset \mathbf{R}^p$ ， $w \in \Omega \subset \mathbf{N}^r$ ，其中 Γ 和 Ω 为解空间的维度， ζ 和 ξ 为从属变量，自变量 $v + \tau$ 的个数表示等式约束的自由度，所以变量 v 和 w 的维度约束应该满足^[10]：

$$v + \tau = n + m - q \quad (16)$$

一般来说 Γ 和 Ω 是任意的， v 和 τ 可以为空，而对于任意一个有效的优化问题，有 $v + \tau \geq 1$ 。假设从属变量 ζ 和 ξ 可以表示为 $\zeta = \zeta(v, w)$ 和 $\xi = \xi(v, w)$ ，优化问题以及约束可重述为^[10]：

$$\begin{cases} f(\mathbf{x}, \mathbf{y}) = f(v, \zeta, w, \xi) = f(v, \zeta(v, w), w, \xi(v, w)) = F(v, w) \\ g(\mathbf{x}, \mathbf{y}) = g(v, \zeta, w, \xi) = g(v, \zeta(v, w), w, \xi(v, w)) = G(v, w) \\ h(\mathbf{x}, \mathbf{y}) = h(v, \zeta, w, \xi) = h(v, \zeta(v, w), w, \xi(v, w)) = H(v, w) \end{cases} \quad (17)$$

优化问题及其约束可转换为以下结构：

$$\begin{cases} \max F(v, w) \\ \text{s.t.} \begin{cases} G(v, w) \geq 0 \\ v \in \Gamma \cap V, w \in \Omega \cap V \\ V = \{(v, w): h(v, \zeta, w, \zeta) = 0, \zeta \in \mathbf{R}^{n-u} \cap X, \zeta \in \mathbf{N}^{m-r} \cap Y\} \end{cases} \end{cases} \quad (18)$$

通过分裂等式约束，将原始变量划分为独立变量和从属变量，原目标优化问题转化为无等式约束问题。从属变量的求解有两种情况：一是从属变量可直接通过等式约束解决；另一种是将独立变量和从属变量的约束使用多目标策略求解，对于从属变量，可将其等式约束的约束视为残差，最小化残差部分的约束即可。将原本含有等式约束的 MINLP 问题转化为只包含不等式约束的优化问题，惩罚函数法结合 PSO 算法就可以求出原本约束范围内的最优解。

2.1.2 算法流程

步骤 1 初始化：

- a) 设置最大迭代次数，速度常量 c_1 、 c_2 以及初始权重；
- b) 随机初始化连续实变量 $v \in \Gamma$ 的粒子位置；
- c) 随机初始化粒子速度；
- d) 根据 sigmoid 方法设置离散粒子 $w \in \Omega$ 的位置；
- e) 开始进入迭代流程。

步骤 2 优化迭代

a) 解出从属变量 ζ 和 ζ 对应的等式约束 $h(v, \zeta, w, \zeta) = 0$ ；寻找可行的 ζ 和 ζ ，以等式 $\sum_{i=1}^q h^2(v, \zeta, w, \zeta)$ 的残差之和最小化为目标；

- b) 评估目标函数 F ；
- c) 更新粒子和群体的最佳值 p_k^i 、 p_k^g 以及粒子的速度向量；
- d) 更新所有粒子的位置；
- e) 满足两个条件之一即刻退出：① 到达最大迭代次数；② $|F^{k+\Delta k} - F^k| < 10^{-5}$, $\Delta k = 15$ ；否则执行下一轮迭代；
- f) 跳转到步骤 2a)；

步骤 3：汇总结果；

步骤 4：结束。

3 仿真与性能分析

根据优化粒子群算法和 SHOT 求解器的数值解法的结果验证本文提出的系统模型以及响应的目标优化问题。该系统在低轨卫星和地面内容服务器以及地面用户的通信网络中共同优化了缓存文件内容的流行度、通信干扰和发射功率的分配以及系统吞吐的对比方案。

SHOT 是一个开源的优化问题求解器^[11]，可单独使用，也可与 GAMS、AMPL、Pyomo 等求解系统^[12]进行集成。本文使用 Pyomo 和 SHOT 求解器进行实验求解。与大多数基于多面体逼近的求解器一样，SHOT 利用对偶策略求解，优化问题的下界由非线性可行集的多面体逼近策略给出^[13]，在此基础上，对原目标优化问题进行重构，求解重构目标的问题，提升原目标优化问题的解。在找到目标问题的近似解后，SHOT 使用边界缩紧在原目标优化问题上寻找更精确的解。具体地说，SHOT 基于区间运算的可行性约束放缩策略来做全局优化，MIP 求解器会自动对问题的线性部分的约束进行收紧^[14]。其算法流程如下简示：

- 1) 自动凸度检测：检测原目标优化问题上的凸度以及非凸的部分；
- 2) 可行性松弛：通过扩展搜索空间，从不可解的子问题的解空间上搜索，将不可解的子问题松弛为可切割的可解问题；
- 3) 自动重构：同步构造原始目标优化问题的对偶子问题和重构问题；
- 4) 整数切割：从搜索空间中排除特定的整数分配，提高解的搜索速度；
- 5) 切割约束：在新形成的约束空间下的解空间中开启搜索全局解。

SHOT 使用多面体逼近策略，在原目标优化问题上检测问题和各项约束的凸度。如果原目标优化问题中存在非凸的区域，标记后使用可行性松弛为凸集。在检测和标记的基础上，扩展搜索空间，从原目标优化问题的不可解的子问题的解空间上搜索，将第一步检测到的子问题松弛为可切割或可支持超平面逼近的可解问题。对整

理后的问题进行重构，和原目标优化问题的对偶问题进行边界缩紧和整数切割，在新的约束空间下找到全局最优解。

3.1 仿真分析

仿真试验中，低轨卫星分别布置 3 颗和 4 颗，地面内容服务器分别布置 3 台和 4 台，地面用户任意。无论缓存是否命中，在 LC 链路和 CG 链路上最大可获得干扰数量为 2 或 3，其他仿真相关参数如表 1 所示。求解目标优化问题分别使用优化的 PSO 算法和 SHOT 求解器。优化的 PSO 算法中，惯性因子为 1，学习因子为 2，种群大小为 200，最大迭代次数为 1 000，粒子的最大速度为 0.5。SHOT 求解器需编译配置使用混合整数规划(Mixed Integer Programming, MIP)求解器和非线性整数规划(Nonlinear Integer Programming, NIP)求解器。

图 2 为信干噪比约束为 8 dB 时，当最大功率约束 $P_{LEO_{max}}$ 在 1~20 W 变化下，分别使用优化的 PSO 算法和 SHOT 求解器在干扰数量分别为 2 和 3 时获得的系统总吞吐。随着功率约束的增加，系统获得的总吞吐不断升高，但系统中增加卫星数目后由于同频干扰的影响，反而降低了系统吞吐，功率约束受 SINR 约束影响，系统吞吐在固定节点下随着功率约束升高而趋于平缓。系统中干扰数量增加给干扰节点的功率带来影响，即需增大通信节点的功率，降低干扰节点的功率才能提高系统吞吐。优化后的 PSO 的算法求解的结果和 SHOT 求解器的结果误差在 1% 之内。在干扰数量为 2 时，系统的总吞吐能达到 45 Mbps 左右，在干扰数量为 3 时，系统的总吞吐能达到 37 Mbps 左右。

图 3 为功率约束为 20 W， $R_{SIN,LC_{min}}$ 在 1~14 dB 变化下，分别使用优化的 PSO 算法和 SHOT 求解器在干扰数量分别为 2 和 3 时获得的系统总吞吐。可以看出，在干扰数量较少时，无论信干噪比在什么情况下，都可以获得较高的系统吞吐，系统中干扰数量每增加一个，对目标优化问题以及约束的影响成几何倍数增加。SINR 约束同样受功率约束影响，随着 SINR 约束增加，通信节点接收到的有用信号强度也随之升高，系统中干扰增加，对通信节点的通信造成压力，从而降低了系统吞吐。具体地，提高系统吞吐也即提高了通信节点的速率，降低了干扰节点的速率，减少了干扰节点同频同时传输对通信节点的影响。

表 1 仿真相关参数

Table 1 Simulation parameters

parameter	value
$d_{LEO,CS_i}/km$	1 200
$d_{CS_i,GU_i}/km$	10
λ/m	0.016 8
B_{LC}/MHz	50
B_{CG}/MHz	20
T_n/K	290
$P_{LEO_{max}}/W$	20
$P_{CS_{max}}/W$	20
δ/s	0.5
F	30
$Q/Mbit$	40
M	20Q
ϑ	0.8
D/m	0.6
$P_{G_{CS_i}}/W$	0.126
$P_{G_{CS_i}'}/W$	0.835
$R_{SIN,LC_{min}}/dB$	8
$R_{SIN,CG_{min}}/dB$	8

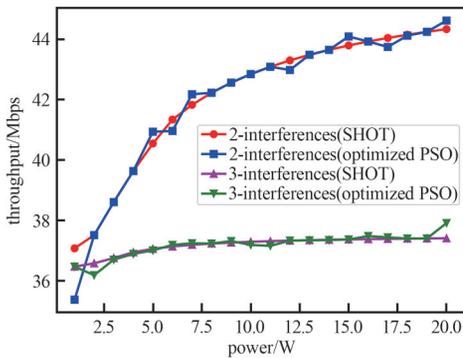


Fig.2 Solution curves for different interference numbers under different power constraints
图 2 不同干扰数在不同功率约束下的求解曲线

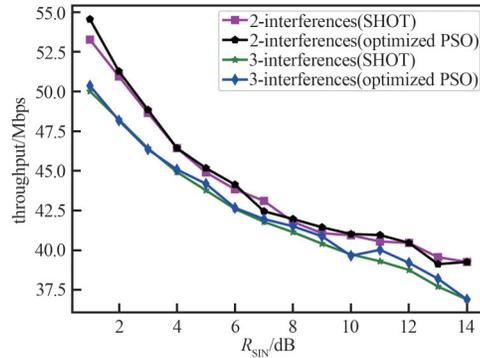


Fig.3 Solution curves for different interference numbers under different SINR constraints
图 3 不同干扰数在不同 SINR 阈值约束下的求解曲线

如图 4 所示，除了通信指标外，缓存是否命中对系统吞吐性能影响较大。图中给出了 $P_{GU_{k,f}}$ 分别为 0.53、0.76、0.98 时，在功率约束为 20 W，信干噪比约束为 8 dB 时的系统吞吐对比。可以看出，缓存命中率越高，系统可达到的吞吐会成倍增加；并且在较高缓存命中率下，系统吞吐的变化率也更高，系统性能提升的也越快。以本次仿真试验为例，系统传输的文件内容在初始状态下包含较多不确定因素，如用户对文件的引用程度。当缓存中某个文件的命中率在短期内迅速提高，系统需面对干扰突变，而系统处于此状态下则满足上文提到的 Zipf 幂律分布。

3.2 算法性能分析

分别设置低轨卫星总数 J 、地面内容服务器总数 I 、地面用户总数 K ，得到不同的目标优化问题，使用优化的 PSO 算法和 SHOT 求解器对不同的目标优化问题进行优化求解。根据表 2 得到的不同目标优化问题求解时间可以看出，在设置低轨卫星总数 $J=4$ 、地面内容服务器 $I=2$ 、地面用户总数 $K=2$ 时，SHOT 求解器求解目标优化问题的运行时间均少于优化的 PSO 算法的运行时间，优化 PSO 算法的运行时间都较为接近。在低轨卫星总数较小时，SHOT 求解器表现较为优秀，但在增加低轨卫星数量以及地面内容服务器数量或地面用户数量后，优化的 PSO 算法运行时间少于 SHOT 求解器的运行时间。在三者数量分别设为 4 后，优化的 PSO 算法运行时间为 19.39 s，SHOT 求解器运行时间为 29.33 s。

从时间复杂度上看，两种算法的运行时间在数据量较小的情况下差别不明显，但在数据量增大后，优化的 PSO 算法运行时间增长速度更慢，时间差距明显，表明优化的 PSO 具有较低的时间复杂度。从空间复杂度上看，由于算法中使用的数据和变量相同，但优化的 PSO 采用了较少的迭代次数，因此在实际应用中可能占用的内存更少。

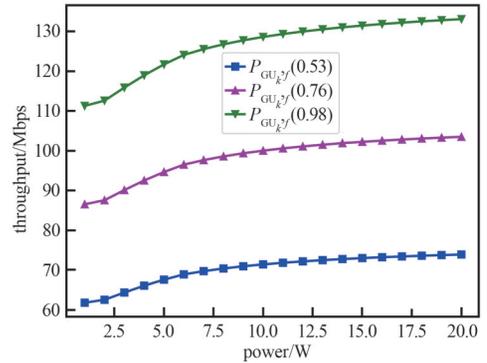


Fig.4 Solution curves under power constraint 20 W for different hit rates
图 4 不同命中率下在功率约束 20 W 下求解曲线

表 2 优化问题求解算法运行时间(单位:s)

Table2 Running time of optimization problem solving algorithm(unit:s)

	optimized PSO algorithm	SHOT solver
$J=2, I=2, K=2$	2.17	0.82
$J=2, I=2, K=4$	2.83	1.56
$J=2, I=4, K=2$	2.97	2.01
$J=4, I=2, K=2$	5.02	6.31
$J=2, I=4, K=4$	7.78	7.62
$J=4, I=2, K=4$	10.08	14.77
$J=4, I=4, K=2$	15.73	22.09
$J=4, I=4, K=4$	19.39	29.33

综上所述，优化的 PSO 算法使用启发式逻辑，在粒子空间中搜索问题的全局最优解，求解受限于粒子空间中粒子的变换，使用空间换时间的思路，在粒子空间中快速迭代寻找问题的解，但在目标问题规模较小时，收益并不明显。相较于 SHOT 求解器，在问题规模较小时，求解速度较快，但在问题规模膨胀之后，为得到相对精确的全局最优解，多面体逼近算法求解问题会浪费大量时间在逼近问题边界的模糊区域。因此，可考虑在较大规模的 MINLP 问题下使用启发式算法求解，在较小规模的问题上使用 SHOT 求解器进行求解。

4 结论

本文研究了低轨卫星和地面内容服务器在缓存辅助下的通信场景，提出了一个系统模型。联合考虑低轨卫星物联网场景下的内容交付网络中低轨卫星、地面内容服务器、缓存以及内容流行度，旨在最大化地面用户可获得的吞吐量。由于综合星地信道模型、地面信道模型以及缓存内容的流行度模型的目标优化问题是一个混合非线性、混合整数非线性规划问题，普通的求解器很难求解。因此本文采用改进的 PSO 算法与新型的 SHOT 求解器对原始目标优化问题进行求解。仿真结果表明，干扰数量的不同，会对系统吞吐产生很大影响，低轨卫星的运行时间和轨迹，以及覆盖地面内容服务器的地面分布，都会在同一频率下对目标传输产生干扰。因此如何在时空环境下优化同频干扰成为下一步研究的方向。同样，在低轨卫星和地面服务器上使用缓存对用户感兴趣的文件进行本地预热，可以减少用户在访问高热度文件的时延，提高系统整体吞吐量。算法分析结果表明，在复杂系统建模下，随着大规模的多目标非线性、非凸非整数规划问题的提出，使用传统数值解法无法满足求解此类问题，而使用启发式算法或使用强化学习等方法为下一步求解复杂问题指出了方向。这项工作将激励未来国内低轨卫星物联网、低轨卫星互联网以及天地一体化等多个方面的相关研究工作，但也将提高对于星地整体系统优化问题的复杂度。本工作也为进一步提高星地网络传输性能指出了下一步的研究重点。

参考文献：

- [1] 吕靖轩,赵国华,邓琳琳. 基于卫星组网的地球观测任务规划与仿真[J]. 火力与指挥控制, 2018,43(11):43-47. (LYU Jingxuan, ZHAO Guohua,DENG Linlin. Earth observation mission planning and simulation based on satellite network[J]. Fire Control & Command Control, 2018,43(11):43-47.)
- [2] ZHANG C,WANG J,WANG L,et al. Research on the security of intelligent transportation systems based on the integration of heaven and earth[J]. Journal of Terahertz Science and Electronic Information Technology, 2021,9(2):97-103.
- [3] LU J,ZHANG J,HE H,et al. Research on Cooperative Network Communication technology based on LEO satellite system[C]// 2019 IEEE the 13th International Conference on Anti-counterfeiting. Xiamen,China:IEEE, 2019:191-195.
- [4] TRAN D H,CHATZINOTAS S,OTTERSTEN B. Satellite-and cache-assisted UAV:a joint cache placement, resource allocation, and trajectory optimization for 6G aerial networks[J]. IEEE Open Journal of Vehicular Technology, 2022(3):40-54. doi:10.1109/OJVT.2022.3142170.
- [5] SHEN Mei,ZHAN Zhihui,CHEN Weineng,et al. Bi-velocity discrete particle swarm optimization and its application to multicast routing problem in communication networks[J]. IEEE Transactions on Industrial Electronics, 2014,61(12): 7141-7151. doi: 10.1109/TIE.2014.2314075.
- [6] LIU Y,LIU X,CAI Y. A hybrid particle swarm optimization algorithm based on opposition-based learning and Gaussian mutation for high-dimensional optimization[J]. Soft Computing, 2019,23(11):4087-4100.
- [7] WANG Z,HE X,HU J. A novel particle swarm optimization algorithm with dynamic population size and multi-parent crossover [J]. IEEE Transactions on Evolutionary Computation, 2020,24(6):1386-1401.
- [8] YANG H,ZHENG W,CHEN M,et al. A hybrid particle swarm optimization algorithm based on opposition-based learning and adaptive learning strategies[J]. IEEE Transactions on Cybernetics, 2021,51(1):284-296.
- [9] LUO X,JIANG J. A Comprehensive survey on swarm intelligence algorithm in multi-objective optimization[J]. IEEE Access, 2021,(9):34990-35005.
- [10] ZHANG L,LIU G,WANG S. A hybrid quantum-inspired PSO algorithm for solving nonconvex MINLP problems[J]. International Journal of Machine Learning and Cybernetics, 2021,12(9):2207-2220.
- [11] LUNDELL A,KRONQVIST J. Polyhedral approximation strategies for nonconvex mixed-integer nonlinear programming in SHOT[J]. Journal of Global Optimization, 2022,82(4):863-896. doi:10.1007/s10898-021-01006-1.
- [12] HART W E,LAIRD C D,WATSON J P,et al. Pyomo—optimization modeling in python[M]. 2nd ed. Berlin,Germany:Springer, 2017.
- [13] BAO Xiaowei,KHAJAVIRAD A,SAHINIDIS N,et al. Global optimization of nonconvex problems with multilinear intermediates[J]. Mathematical Programming Computation, 2015,7(1):1-37. doi:10.1007/s12532-014-0073-z.
- [14] LUNDELL A,KRONQVIST J,WESTERLUND T. The supporting hyperplane optimization toolkit for convex MINLP[J]. Journal of Global Optimization, 2022,84(1):1-41. doi:10.1007/s10898-022-01128-0.

作者简介：

陆涣冰(1997-), 男, 在读硕士研究生, 主要研究方向为卫星通信 .email:luhuanbing084@gmail.com.

任壮壮(1997-), 男, 在读硕士研究生, 主要研究方向为卫星通信.

丁晓进(1981-), 男, 博士, 教授, 主要研究方向为卫星通信、空间信息网络、卫星物联网.

张更新(1967-), 男, 博士, 教授, 主要研究方向为天地一体化网络、卫星物联网.