

机器类通信中基于NOMA短编码块传输的高可靠低延迟 无线资源分配优化方案

谢显中 黎佳* 黄倩 陈杰

(重庆邮电大学通信与信息工程学院计算机网络与通信重庆市重点实验室 重庆 400065)

摘要: 针对机器类通信(MTC)应用场景的业务特征和服务质量(QoS)要求, 该文考虑基于非正交多址(NOMA)的MTC中短分组/短编码块传输, 探讨MTC中基于NOMA的高可靠低延迟无线资源优化问题。首先, 上行传输是基于NOMA的MTC通信的瓶颈, 考虑无线蜂窝网络中支持NOMA和高可靠低延迟性能要求, 该文建立了上行无线资源优化的系统模型; 然后, 分析上行传输延迟, 导出基于距离的链路可靠性函数; 进一步, 以延迟、可靠性和带宽为约束下条件, 提出一种最大化中心用户和速率的无线资源分配算法, 并给出算法的收敛性证明和复杂度分析; 最后, 实验仿真验证了所提算法的性能优势。

关键词: 机器类通信; 高可靠低延迟; 短编码块传输; 非正交多址; 无线资源优化

中图分类号: TN925

文献标识码: A

文章编号: 1009-5896(2019)11-2549-08

DOI: [10.11999/JEIT190128](https://doi.org/10.11999/JEIT190128)

Optimal Scheme of Resource Allocation for Ultra-reliable and Low-latency in Machine Type Communications Based on Non-orthogonal Multiple Access with Short Block Transmission

XIE Xianzhong LI Jia HUANG Qian CHEN Jie

(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications Chongqing key Laboratory of Computer Networks and Communications, Chongqing 400065, China)

Abstract: For the service characteristics and Quality of Service (QoS) requirements of Machine Type Communications (MTC), short-packet/short-coded block transmission in MTC based on Non-Orthogonal Multiple Access (NOMA) is considered in this paper, and the resource optimization problem of the Ultra-Reliable and Low-Latency (URLL) in MTC based on NOMA is discussed. Currently, uplink transmission is a bottleneck of MTC based on NOMA. Firstly, considering the performance requirements supporting NOMA and high reliability and low latency in wireless cellular networks, a system model for uplink wireless resource optimization is established. Then, the uplink transmission delay is analyzed and the link reliability function based on distance is derived. Further, with the constraints of delay, reliability and bandwidth, a wireless resource allocation algorithm for maximizing the sum rates of central users is proposed, and also the convergence proof and complexity analysis of the algorithm are given. Finally, the simulation results show the performance advantages of the proposed optimal scheme.

Key words: Machine Type Communications(MTC); Ultra-Reliable and Low-Latency(URLL); Short block transmission; Non-Orthogonal Multiple Access(NOMA); Resource optimization

收稿日期: 2019-03-05; 改回日期: 2019-06-26; 网络出版: 2019-07-04

*通信作者: 黎佳 li_jia_2017@163.com

基金项目: 国家自然科学基金(61601070), 重庆市教委科学技术研究重点项目(KJZD-K201800603), 重庆市基础与前沿研究计划项目(cstc2018jcyjAX0432), 重庆市研究生科研创新项目(CYS17223)

Foundation Items: The National Nature Science Foundation of China (61601070), The Key Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-K201800603), The Foundation and Frontier Research Program of Chongqing (cstc2018jcyjAX0432), The Graduate Scientific Research Innovation Program of Chongqing (CYS17223)

1 引言

面向2020年及未来, 机器类通信(Machine Type Communications, MTC)将带来越来越多样化的需求, 以及移动通信业务量和资源需求的剧增^[1]. MTC并发性地产生大量的短数据包类业务请求, 而且部分应用场景还要求在支持大连接数的同时, 进一步提供高可靠低延迟(Ultra-Reliable and Low-Latency, URLL)的服务质量(Quality of Service, QoS)^[2].

综合5G中MTC场景的海量移动用户数量及其业务特征, 一些文献研究了不同类型的无线接入技术^[3], 但是传统正交资源分配方式显然限制了额外用户数的接入. 非正交多址(Non-Orthogonal Multiple Access, NOMA)技术能够通过功率分配和串行干扰消除(Successive Interference Cancellation, SIC)技术使得多个用户共享同一资源, 实现时/频域上的非正交资源分配, 从而提升系统的频谱效率和能量效率, 并显著增加连接设备的数量^[4]. 对于这种功率域NOMA, 信道质量差的用户可以分到更多的功率, 以确保达到用户的目标速率, 从而在网络吞吐量和用户公平性之间取得平衡^[5]. 因此, 将MTC与NOMA技术相结合, 可以更好地提升MTC中用户服务质量.

文献^[6,7]讨论了提供低延迟高可靠(URLL)的MTC, 文献^[8,9]等讨论了提供URLL的NOMA. 但是, 文献^[6,7]未涉及NOMA接入方式, 文献^[8,9]未涉及MTC通信问题. 相对于满足人类通信需求的传统H2H通信, MTC的无线资源分配与传统实时业务的核心区别在于传输延迟和数据包大小, 一般传输延迟更小(比如1 ms左右), 采用短数据包(比如20 Byte)传输^[10], 而H2H通信的传输延迟相对较长(如10 ms)和数据包大小较大(如1500 Byte)^[11]. 因此, MTC中信道编码的块长较短.

文献^[12-16]研究了MTC的无线资源分配方案. 文献^[12]提出了在LTE网络M2M/H2H共存场景中, 在QoS的保证下上行链路资源分配设计, 利用正则化对偶理论(CDT)解决了资源分配问题. 文献^[13]将联合M2M和H2H通信调度作为优化问题, 该问题的最优解最大限度地提高了H2H的吞吐量, 同时满足了对延迟敏感的MTC设备延迟要求. 文献^[14]分析了URLLC网络中的延迟和丢包率相关要求和功能实体, 以及支持用户服务质量的网络可用性, 并对满足URLLC要求的网络资源优化进行了讨论, 但没有给出具体的方案. 文献^[15,16]研究了URLLC通信模型的资源分配问题.

但是, 上述文献大多数是在信道编码无限长的信道中研究, 没有考虑在低时延约束要求. MTC

的无线资源分配与传统实时业务的核心区别在于传输延迟和数据包大小, 在信道编码的长度有限长时, 用户的解码错误概率需考虑在内; 此外, 如果将信道编码无限长信道的香农容量用于设计资源分配, 则不能保证排队时延和时延违反概率, 从而不能满足高可靠性.

基于以上分析, 本文探讨MTC中基于NOMA的高可靠低延迟无线资源优化问题. 首先, 考虑无线蜂窝网络中支持NOMA和高可靠低延迟性能要求, 建立上行无线资源优化的系统模型; 然后, 分析上行传输延迟, 导出基于距离的链路可靠性函数; 进一步, 以延迟、可靠性和带宽为约束条件, 提出一种最大化中心用户和速率的无线资源分配算法, 并给出收敛性证明和复杂度分析; 最后, 实验仿真验证了所提算法的性能优势.

2 系统模型

本文考虑非正交多址系统单小区上行链路, 如图1所示, 1个基站(Base Station, BS)的蜂窝系统, 蜂窝基站天线的数目为 N_t , K 对单天线MTC用户(含2个MTCD), 即系统中MTCD的总数为 $2K$, 1个用户对共用同一信道, 假设用户在发送传输请求前都需要接入到基站中, 用户对产生的数据发送给基站, 基站成功接收数据后发送给接收端. 由于多天时分双工(Time Division Duplexing, TDD)基站有测距(用户到基站的距离)能力, 本文采用TDD方式.

本文考虑频率-平坦衰落信道, 当用户设备有数据包传输时才分配带宽, 基站给第 k 对用户分配 N_k 个子信道, B_k 表示分配给第 k 个用户对每个子信道的带宽, W_c 表示信道相干带宽, 当分配给每个用户对带宽小于信道相干带宽时它是有效的, 即 $B_k < W_c$ 能满足此要求, 分配给第 K 个用户对带宽为 $N_k B_k$, 每个子信道具有独立的信道增益, 并发传输的用户对设备使用不同的子信道.

NOMA是基于功率域分配的新型多址技术, 它通过将多个用户信号在功率域进行简单线性叠加, 实现在相同的时域、频域或空域资源上的传输. 因为并发传输的用户对设备使用不同的子信道, 本文只考虑1个用户对之间的干扰, 下面举例说明. 在图1中, 假设现有用户1和用户2占用相同的时域、频域、空域资源, 两者的信号在功率域上叠加, 用户1是小区的边缘用户, 用户2是小区的中心用户. 其中用户1分配较高的功率, 用户2分配较低的功率. 在接收端, 假设基站使用SIC接收机来解码多用户信号, 用户2首先进行SIC, 将用户1的信号从接收信号中减去, 随后解码自己的信号, 此

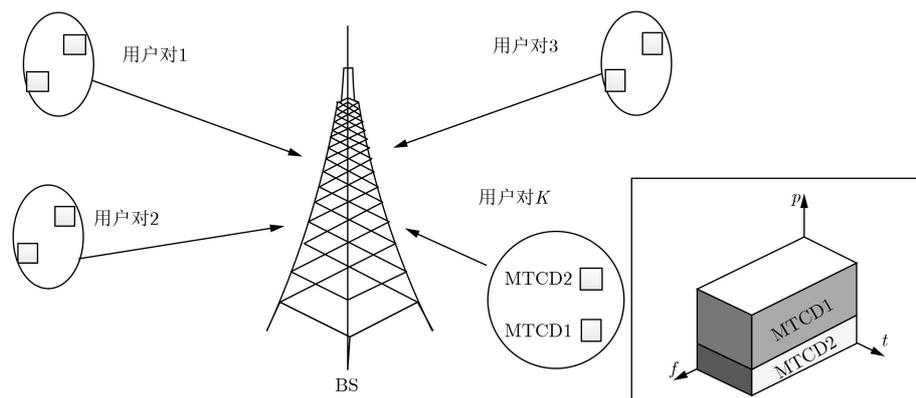


图1 非正交多址K个用户对系统模型

时, 用户2不受用户1的干扰。 P_k^1 表示边缘用户的发射功率, P_k^2 表示中心用户的发射功率, 由于多天线的TDD基站具有测距能力, 在TDD上行传输中采用反向功率控制。

用户设备的发射功率 P_k^i ($i = 1, 2$)表示为

$$P_k^i = \begin{cases} \rho_0 r_k^\alpha, & \rho_0 r_k^\alpha < P_{\max} \\ P_{\max}, & \text{其他} \end{cases} \quad (1)$$

其中, r_k 为用户离基站的距离, $\alpha > 2$ 为路径损耗指数, ρ_0 为基站所需的平均接收功率, P_{\max} 为用户的最大传输功率。

基站接收到第 k 个用户对中第 i 个用户的信号为

$$y_k^i = h_k^i \left(\sqrt{P_k^1} x_k^1 + \sqrt{P_k^2} x_k^2 \right) + n_k^i \quad (2)$$

其中, 信道系数 $h_k^i \sim \text{CN}(0, 1) \in \mathbb{C}^{N_t \times 1}$, 即服从独立同分布的高斯分布, 其均值为0, 方差为1。 x_k^i 为第 k 个用户对中 i 个用户发送的信号, n_k^i 为加性高斯白噪声, 其均值为0, 方差为 σ_n^2 。

为了满足系统的延迟需求, 由于不能采用重传机制, 本文考虑一个基于可实现速率的传输模型。第 k 个用户对在第 j ($j \in (0, N_k)$)个子信道上的和速率表示为 $R_{k,j}^i$, 有

$$R_{k,j}^1 = \frac{1}{\ln 2} \left\{ \ln \left[1 + \frac{\mu_k^1 P_k^1 g_{k,j}^1}{N_0 B_k + \mu_k^1 P_k^2 g_{k,j}^1} \right] - \sqrt{\frac{V_{k,j}^1}{T_f B_k}} f_Q^{-1}(\epsilon_{k,j}^1) \right\} \quad (3)$$

$$R_{k,j}^2 = \frac{1}{\ln 2} \left\{ \ln \left[1 + \frac{\mu_k^2 P_k^2 g_{k,j}^2}{N_0 B_k} \right] - \sqrt{\frac{V_{k,j}^2}{T_f B_k}} f_Q^{-1}(\epsilon_{k,j}^2) \right\} \quad (4)$$

其中, 每个TDD帧的持续时间为 T_f , 等同于系统的

传输时间间隔 (Transmission Time Interval, TTI), 第 k 个用户对的平均信道增益表示为 μ_k^i , 第 k 个用户对在第 j 个子信道上的瞬时信道增益表示为 $g_{k,j}^i = (h_{k,j}^i)^H h_{k,j}^i$, $[\cdot]^H$ 表示共轭转置。 N_0 表示单边信号噪声谱密度, $f_Q^{-1}(x)$ 表示Q函数的逆函数, $V_{k,j}^1 = 1 - [1 + \mu_k^1 P_k^1 g_{k,j}^1 / (N_0 B_k + \mu_k^1 P_k^2 g_{k,j}^1)]^{-2}$, $V_{k,j}^2 = 1 - [1 + \mu_k^2 P_k^2 g_{k,j}^2 / N_0 B_k]^{-2}$ 。

3 传输延迟与可靠性分析

为提供URLL的QoS和资源分配策略, 针对每个用户对的数据包的上传传输延迟和总体可靠性要求, 考虑短帧结构系统, 本文得到了上行传输延迟的表达式和基于距离的可靠性函数, 下面分别讨论。

3.1 上行传输延迟

为了满足系统的低延迟需求, 不采用重传机制, 第 K 个用户对的数据包的QoS需求定义为第 k 个用户对由控制信号和数据传输引起的延迟 D_k^i 、第 k 个用户对的数据传输引起的延迟 ϵ_k^i , 其中 $k=1, 2, \dots, K$ 。 D_{\max} 表示可允许最大的端到端延迟, ϵ_{\max} 表示可允许最大的传输错误率, 为了满足QoS需求, 第 k 个用户对必须满足 $D_k^i < D_{\max}$ 和 $\epsilon_k^i < \epsilon_{\max}$ 。为了减少传输延迟, 本文考虑用户数相对稳定的情况, 借鉴文献[17]中提出的短帧结构, 如图2所示, 假设忽略处理延迟, 端到端延迟包括上行传输延迟 D_u 和下行传输延迟 D_d 、在基站上排队延迟 D_q 和回程延迟 D_b , 所以在上行链路由控制信号和数据传输引起的延迟为 $D_k^i = 4T_f$ 。

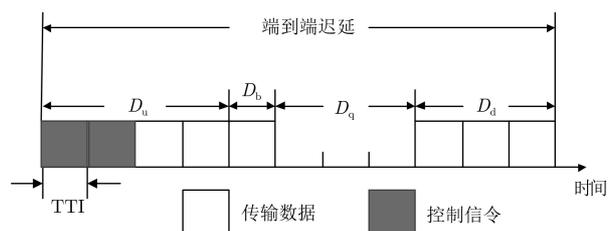


图2 短帧结构

3.2 基于距离的可靠性函数

为了满足系统的延迟需求, 由于不能采用重传机制, 本文考虑一个基于可实现速率的传输模型。设每个数据包的大小为 u bit, 根据香农容量公式的无限长度编码, 当基站完全知道 μ_k^i 和 $h_{k,j}^i$ 的信息时, 第 k 个用户对到基站在一帧内能传输的最大比特数为

$$s_{k,j}^1 = \frac{T_f B_k}{\ln 2} \left\{ \ln \left[1 + \frac{\mu_k^1 P_k^1 g_{k,j}^1}{N_0 B_k + \mu_k^1 P_k^2 g_{k,j}^1} \right] - \sqrt{\frac{V_{k,j}^1}{T_f B_k}} f_Q^{-1}(\varepsilon_{k,j}^1) \right\} \quad (5)$$

$$s_{k,j}^2 = \frac{T_f B_k}{\ln 2} \left\{ \ln \left[1 + \frac{\mu_k^2 P_k^2 g_{k,j}^2}{N_0 B_k} \right] - \sqrt{\frac{V_{k,j}^2}{T_f B_k}} f_Q^{-1}(\varepsilon_{k,j}^2) \right\} \quad (6)$$

该模型基于有限块长度信道编码, 如果接收到的信噪比高于门限值, 则可以成功地传输 u bit 的数据包, 否则, 传输则不能成功。因此, 如果瞬时信道增益 $g_{k,j}^i$ 高于门限值 $g_k^{i,\text{th}}$, 则该数据包可以以 $1 - \varepsilon_{k,j}^i$ 的概率成功传输。否则, 系统就丢掉数据包。因为 $f_Q^{-1}(\cdot)$ 是一个减函数, 所以 $g_{k,j}^i$ 在 $\varepsilon_{k,j}^i$ 上是减函数, 有

$$\varepsilon_{k,j}^i \leq \begin{cases} \varepsilon_k^{i,\text{th}}, & g_{k,j}^i \geq g_k^{i,\text{th}} \\ 1, & g_{k,j}^i < g_k^{i,\text{th}} \end{cases} \quad (7)$$

在很高的信噪比下 $V_{k,j}^i \approx 1$, 把 $s_{k,j}^i = u$, $V_{k,j}^i = 1$ 代入得到

$$g_k^{1,\text{th}} = \frac{N_0 B_k}{\mu_k^1} \frac{\exp \left[\frac{u \ln 2}{T_f B_k} + \sqrt{\frac{1}{T_f B_k}} f_Q^{-1}(\varepsilon_k^{1,\text{th}}) \right] - 1}{P_k^1 - P_k^2 \left\{ \exp \left[\frac{u \ln 2}{T_f B_k} + \sqrt{\frac{1}{T_f B_k}} f_Q^{-1}(\varepsilon_k^{1,\text{th}}) \right] - 1 \right\}} \quad (8)$$

$$g_k^{2,\text{th}} = \frac{N_0 B_k}{\mu_k^2 P_k^2} \left\{ \exp \left[\frac{u \ln 2}{T_f B_k} + \sqrt{\frac{1}{T_f B_k}} f_Q^{-1}(\varepsilon_k^{2,\text{th}}) \right] - 1 \right\} \quad (9)$$

如果第 k 个用户对成功传输, 瞬时信道增益的门限值 $g_k^{\text{th}} = \max(g_k^{1,\text{th}}, g_k^{2,\text{th}})$ 。

本文的传输模型考虑了当 $g_{k,j}^i \geq g_k^{\text{th}}$ 时的传输错误概率, 因此适用于在超高可靠性约束下设计的传输策略。第 k 个用户对中的用户1能成功传输的概率是至少有1个子信道的瞬时信道增益大于或等于 g_k^{th} 即

$$\Pr \left\{ \bigcup_{j=1}^{N_k} \{g_{k,j}^1 \geq g_k^{\text{th}}\} \right\} = 1 - \prod_{j=1}^{N_k} \Pr \{g_{k,j}^1 < g_k^{\text{th}}\} = 1 - \left[\int_0^{g_k^{\text{th}}} f_g(x) dx \right]^{N_k} \quad (10)$$

第 k 个用户对中的用户2能成功传输的概率是至少有1个子信道的瞬时信道增益大于或等于 g_k^{th} 即

$$\Pr \left\{ \bigcup_{j=1}^{N_k} \{g_{k,j}^2 \geq g_k^{\text{th}}\} \right\} = 1 - \prod_{j=1}^{N_k} \Pr \{g_{k,j}^2 < g_k^{\text{th}}\} = 1 - \left[\int_0^{g_k^{\text{th}}} f_g(x) dx \right]^{N_k} \quad (11)$$

其中, $f_g(x) = \frac{1}{(N_t - 1)!} x^{N_t - 1} e^{-x}$ 表示瞬时信道增益分布函数。

用 $1_{k,j} = 1$ 表示第 k 个用户对能成功传输, $1_{k,j} = 0$ 表示不能成功传输, 所以, 第 k 个用户对能成功传输的概率为

$$\Pr \{1_{k,j} = 1\} = \Pr \left\{ \bigcup_{j=1}^{N_k} \{g_{k,j}^1 \geq g_k^{\text{th}}\} \right\} (1 - \varepsilon_{k,j}^1) \cdot \Pr \left\{ \bigcup_{j=1}^{N_k} \{g_{k,j}^2 \geq g_k^{\text{th}}\} \right\} (1 - \varepsilon_{k,j}^2) \approx 1 - 2 \left[\int_0^{g_k^{\text{th}}} f_g(x) dx \right]^{N_k} - \varepsilon_{k,j}^1 - \varepsilon_{k,j}^2 \quad (12)$$

其中, $\varepsilon_{k,j}^1$, $\varepsilon_{k,j}^2$ 和 $\left[\int_0^{g_k^{\text{th}}} f_g(x) dx \right]^{N_k}$ 的值都极小, 其他值可以忽略。

从式(12)可知, 影响链路可靠性的参数主要包括 N_k , $\varepsilon_{k,j}^1$, $\varepsilon_{k,j}^2$, B_k 及 r_k 。设 $f_k(r_k, N_k, B_k, \varepsilon_{k,j}^1, \varepsilon_{k,j}^2) \triangleq \left[\int_0^{g_k^{\text{th}}} f_g(x) dx \right]^{N_k} + \frac{1}{2} (\varepsilon_{k,j}^1 + \varepsilon_{k,j}^2)$ 。其中, $f_k(r_k, N_k, B_k, \varepsilon_{k,j}^1, \varepsilon_{k,j}^2)$ 包括距离因素, 且距离因素在后面方案中起重要作用, 称为基于距离的可靠性函数, 于是有

$$\Pr \{1_k = 1\} \approx 1 - f_k(r_k, N_k, B_k, \varepsilon_{k,j}^1, \varepsilon_{k,j}^2) \quad (13)$$

4 最大化中心用户和速率的无线资源优化算法

在确保用户对达到了延迟和可靠性的性能前提下，最大化中心用户(用户2)系统的和速率为本文的优化目标。

在延迟、可靠性和带宽约束下，最大化中心用户和速率的无线资源优化问题可以描述为

$$\max_{\Delta} \sum_{k=1}^K \sum_{j=1}^{N_k} R_{k,j}^2 \quad (14)$$

$$\text{s.t.} \quad 0 < D_k^j \leq D_{\max} \quad (15)$$

$$0 < \varepsilon_{k,j}^i \leq \varepsilon_{\max} \quad (16)$$

$$f_k(r_k, N_k, B_k, \varepsilon_{k,j}^1, \varepsilon_{k,j}^2) \leq \varepsilon_{\max} \quad (17)$$

$$0 < B_k \leq W_c \quad (18)$$

$$0 < N_k \leq N_{\max} \quad (19)$$

其中， $\Delta = \{N_k, B_k, \varepsilon_{k,j}^1, \varepsilon_{k,j}^2\}$ 为参数集，式(15)为延迟约束，式(16)和式(17)为可靠性约束，式(18)和式(19)表示带宽约束。

一般系统中的帧长是确定的，参考文献[13]，本文设置每个帧的持续时间 $T_f=0.1$ ms。此外，由于每个用户对约束条件不受其他用户影响，式(14)中的问题可以等价的分解为单用户对设备问题，即

$$\max_{\Delta} \sum_{j=1}^{N_k} R_{k,j}^2 \quad (20)$$

s.t. 式(15)—式(19)

定理1 当 $B_k \geq 1$ 时， $R_{k,j}^2$ 在 B_k 上是减函数。

证明 在高信噪比下， $\mu_{k,j}^2 P_k^2 g_{k,j}^2 / (N_0 B_k) \gg 1$ ， $V_{k,j}^2 \approx 1$ ，当 $\varepsilon_{k,j}^2 \in [10^{-7}, 10^{-5}]$ ，查表得到 $f_Q^{-1}(\varepsilon_{k,j}^2) \in [4.2649, 5.1995]$ ，本文设定 $T_f = 0.1$ ms = 10^{-4} s。

令 $A = \mu_{k,j}^2 P_k^2 g_{k,j}^2 / N_0$ ， $C = f_Q^{-1}(\varepsilon_{k,j}^2) / \sqrt{T_f} \gg 1$ ，则对 $R_{k,j}^2$ 求导得

$$\begin{aligned} \frac{\partial R_{k,j}^2}{\partial B_k} &= -\frac{A}{B_k^2 + AB_k} + \frac{C}{2} B_k^{-\frac{3}{2}} \\ &= \frac{\frac{C}{2} B_k^{\frac{1}{2}} + \frac{AC}{2} B_k^{-\frac{1}{2}} - A}{B_k^2 + AB_k} \\ &= \frac{\frac{C}{2} \left(B_k^{\frac{1}{2}} + AB_k^{-\frac{1}{2}} \right) - A}{B_k^2 + AB_k} \end{aligned} \quad (21)$$

由不等式性质得到，当 $a \geq 1, b \geq 1$ 时有 $a + b \leq 2ab$ 。

$$AB_k^{-\frac{1}{2}} = \frac{\mu_{k,j}^2 P_k^2 g_{k,j}^2}{N_0 B_k} \times \sqrt{B_k} \geq 1 \quad (22)$$

$$\begin{aligned} \frac{\partial R_{k,j}^2}{\partial B_k} &= \frac{\frac{C}{2} \left(B_k^{\frac{1}{2}} + AB_k^{-\frac{1}{2}} \right) - A}{B_k^2 + AB_k} \leq \frac{CA - A}{B_k^2 + AB_k} \\ &= \frac{A(C - 1)}{B_k^2 + AB_k} < 0 \end{aligned} \quad (23)$$

所以， $\frac{\partial R_{k,j}^2}{\partial B_k} < 0$ ，即 $R_{k,j}^2$ 在 B_k 为减函数，即当 B_k 最小时， $R_{k,j}^2$ 有最优解。证毕

第 k 用户对中的用户2的和速率为

$$\begin{aligned} R_k^2 &= \sum_{j=1}^{N_k} R_{k,j}^2 = \sum_{j=1}^{N_k} \frac{1}{\ln 2} \left\{ \ln \left[1 + \frac{\mu_{k,j}^2 P_k^2 g_{k,j}^2}{N_0 B_k} \right] \right. \\ &\quad \left. - \sqrt{\frac{V_{k,j}^2}{T_f B_k}} f_Q^{-1}(\varepsilon_{k,j}^2) \right\} \end{aligned} \quad (24)$$

可以得到 R_k^2 的值与第 k 用户对所需的总带宽($N_k B_k$)有关，由定理1可得 $R_k^2 R_k^2$ 与带宽成反比例。

因为用户数的数量很大，而所需的数据率很低(短数据包传输)，在实际系统中可把频率离散成基本的带宽单元，每个子信道由多个带宽单元组成，通过调整子信道上的带宽单元数，可以改变每个子信道的带宽，从而使用户2的和速率最大化。

基于上述讨论，本文提出的延迟、可靠性和带宽约束下最大化中心用户和速率的无线资源分配算法如表1。

根据定理1算法是收敛的。对于复杂性，算法的复杂度与用户2的和速率表达式(4)有关，即与最大子信道个数 N_{\max} 有关，其复杂度直接由迭代次数 N_{\max} 决定，令式(4)的运算量为 $O(1)$ ，则算法1的运行复杂度为 $O(N_{\max})$ ，为了解决式(14)的无线资源优化问题，由于每个用户对约束条件不受其他用户影响，需要对式(20)进行 K 次求解，所以本文所提的算法复杂度为 $O(K \times N_{\max})$ ，其复杂度的高低取决于用户数 K 和最大子信道个数 N_{\max} 的取值。由此可见，本文所提出的最小化系统所需带宽的分配算法计算复杂度随着用户数量的增加而成线性增加，能较大程度的减小整个方案的复杂度。

5 仿真分析

利用Matlab软件仿真验证本文算法的性能。系统包括1个蜂窝基站、MTC用户对数为 $K=500$ ，路径损耗模型为 $-10 \lg(\mu_k^i) = 35.3 + 37.6 \lg(r_k)$ ，用户在离基站 $[50, d]$ m之间均匀分布， d 是小区的半径，因为平均信道增益 μ_k^i 是随着 r_k 递减的，用户离基站的距离越远，需要更多的资源来保证用户的

表1 算法1的具体流程

算法1 用户2的和速率最大化算法	
步骤1	输入 N_k, r_k , 当 $N_k \leq N_{\max}$ 时, 初始值 $B_{lb} = 0, B_{ub} = W_c, B_0 = \frac{B_{lb} + B_{ub}}{2}$;
步骤2	当 $B_{ub} - B_{lb} > \delta_b$, 通过二分法计算得到使 $f_u(r_k, N_k, B_i, \epsilon_i^1, \epsilon_i^2)$ 最小化的 $\epsilon_i^1, \epsilon_i^2$;
步骤3	若 $f_u(r_k, N_k, B_i, \epsilon_i^1, \epsilon_i^2) > \epsilon^{\max}$, 更新初始值 $B_{lb} = B_i, B_i = \frac{B_{lb} + B_{ub}}{2}$, 或者 $B_{ub} = B_i, B_i = \frac{B_{lb} + B_{ub}}{2}$; 否则转步骤4;
步骤4	若 $f_u(r_k, N_k, B_i, \epsilon_i^1, \epsilon_i^2) < \epsilon^{\max}, B_k(N_k) = B^{\text{opt}}$ 或者 $B_k(N_k) = \text{NaN}$;
步骤5	求出 $N_k^* = \arg \min_{N_k} N_k B_k(N_k), B_k^* = B_k(N_k^*)$;
步骤6	把 N_k^*, B_k^* 代入到 R_k^2 中求出最优的 R_k^{2*} .

服务质量, 所以考虑 $r_k > 50 \text{ m}$, 用户的最大发射功率 $P_{\max} = 23 \text{ dBm}$, 基站所需的平均接收功率 $\rho_0 = -40 \text{ dBm}$, 路径损耗指数 $\alpha = 3$, 包长20 Byte, $N_0 = -174 \text{ dBm/Hz}$, $W_c = 0.5 \text{ MHz}$, $T_f = 0.1 \text{ ms}$, 给用户分配的最大子信道数 $N_{\max} = 10$.

MTC设备应用场景中, 对QoS需求中可靠性性能要求很高, 图3给出了当 $\epsilon_{\max} = 10^{-5}$ 时, 用户的可靠性随着用户离基站的距离变化的关系曲线, 从图3可以看出, 随着距离的增大, 用户的可靠性降低, 这是因为离基站越远, 信道的质量会降低, 随之可靠性也会有所降低, 但用户的可靠性性能曲线都在99.999%以上, 所有采用本文方案可以满足MTC对QoS中可靠性性能的需求。当基站的天线数增加时, 用户的可靠性会有所增加, 这得益于天线数的增加带来的增益。

在不同的MTC设备应用场景中, 对QoS需求中可靠性性能要求可能不相同, 因此, 图4和给出了当 $\epsilon_{\max} = 10^{-5}, \epsilon_{\max} = 10^{-6}$ 和 $\epsilon_{\max} = 10^{-7}$ 时, 天线数为8, 16和32情况下系统和速率与用户数的关系曲线。从图4以看出, 随着天线数的增大, 不同 ϵ_{\max} 下, 系统的和速率均增大, 这得益于天线数的增加带来的增益。当对可靠性的要求提高时, 用户对分配到的子信道带宽会有所增加, 这是因为当系统的QoS需求中可靠性性能提高时, 对传输信道的质量要求更高, 系统的和速率会有所减少。

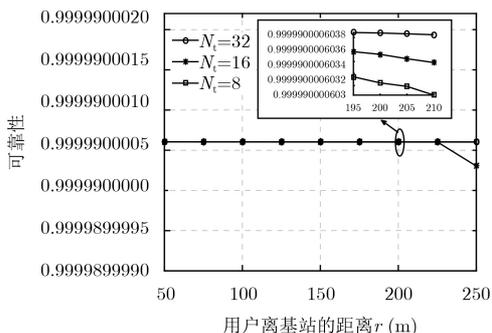


图3 用户离基站的距离与可靠性的关系曲线

在仿真中, 本文将文献[14]中采用的OMA方案做为对比方案。图5给出了不同方案中系统需要的总带宽与天线数的关系曲线, $\epsilon_{\max} = 10^{-5}$, 从图5中可以看出, 随着天线数的增加, 系统所需带宽随之减少, 在NOMA方案中, 当天线数为32时, 所需要的带宽大约为82 MHz, 而当天线数为8时, 所需要的带宽大约为130 MHz, 这是因为随着天线数的增加, 增加了空域的多样性, 用户分配得到的子信道数会减少, 这使得系统占用的总带宽减少。对比OMA方案, 两者在相同天线数和MTC总用户数相同情况下, 系统占用的总带宽大约增加了40%, 本文NOMA方案明显优于文献[14]中OMA方案。

图6为不同方案下系统吞吐量随用户数变化的关系曲线, 由图可知, 当天线数增加时, 两个方案

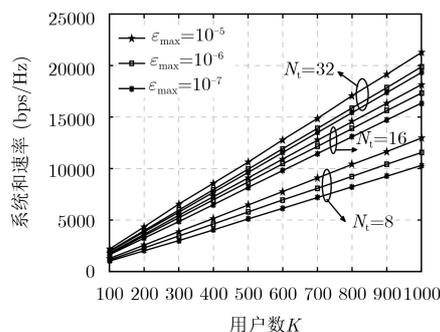


图4 不同 ϵ_{\max} 下用户数与系统和速率的关系曲线

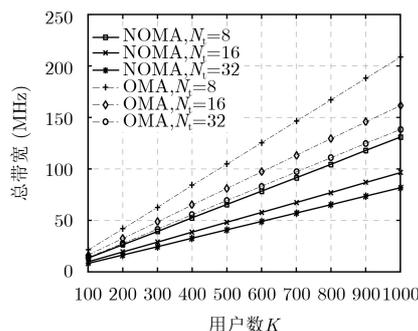


图5 用户数与系统所需总带宽的关系曲线

下的总吞吐量都会增加，这是因为当天线数增加时，增加了空域的多样性，从而增加了总吞吐量；另外，与OMA方案相比较，本文方案的吞吐量性能优于OMA方案，从图6可以看出，两者在相同天线数和MTC总用户数相同情况下，NOMA方案中系统吞吐量大于OMA方案中系统吞吐量，但是，从图5可以看出，NOMA方案中系统所需带宽明显小于OMA方案中系统所需带宽，因此，也可以得出NOMA方案中的和速率性能也优于文献[14]中OMA方案中的和速率性能。

图7为不同方案下能量效率随用户数变化的关系曲线，能量效率定义为系统的和速率与总发射功率之间的比值，由图7可知，在不同天线数下，能量有效性曲线趋势基本保持不变，这是因为当用户数增加时，系统的和速率和总功率也会随之增加，他们的比值所以基本保持不变；当天线数增加时，由图4可知，系统的和速率会增大，而用户的发射功率与天线数无关，所以天线数多的能量效率性能比同方案下天线数少的性能要好；另外，与文献[14]中所提的OMA方案相比较，本文方案的能量效率明显优于文献[14]方案，这是因为本文方案的系统和速率性能优于文献[14]方案的系统和速率性能，因此得到的能量效率性能比文献[14]方案中的能量效率性能好。

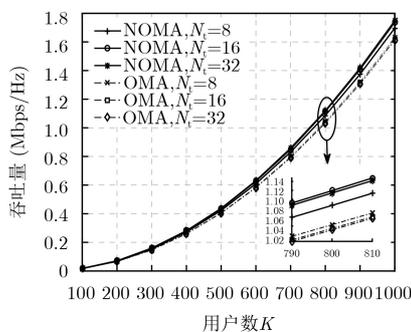


图6 用户数与吞吐量的关系曲线

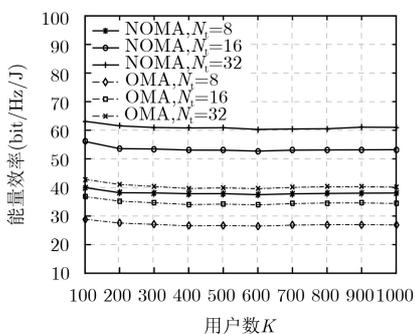


图7 用户数与能量效率的关系曲线

6 结束语

本文研究了基于NOMA的MTC中上行短分组/

短编码块传输的无线资源优化，提出一种高可靠低延迟无线资源优化方案。当MTC场景中移动用户量激增时，为在带宽受限下满足高可靠低延迟要求，考虑一个基于可实现速率的传输模型。基于有限块长度信道编码的可实现速率来反映可靠性约束，采用NOMA技术，实现时/频域上的非正交资源分配，满足系统QoS需求的接入和传输，并进行了收敛和复杂度分析。仿真验证了本文方案在系统所需带宽、吞吐量、能量效率等性能方面都有一定的优势。下一步，将结合该场景的实际应用，针对海量用户的过载接入问题，实现带宽受限下高可靠低延迟通信方案。

参考文献

- [1] WANG Mao, YANG Wenjie, ZOU Jun, *et al.* Cellular machine-type communications: Physical challenges and solutions[J]. *IEEE Wireless Communications*, 2016, 23(2): 126–135. doi: [10.1109/MWC.2016.7462494](https://doi.org/10.1109/MWC.2016.7462494).
- [2] BENNIS M, DEBBAH M, and POOR H V. Ultrareliable and low-latency wireless communication: tail, risk, and scale[J]. *Proceedings of the IEEE*, 2018, 106(10): 1834–1853. doi: [10.1109/JPROC.2018.2867029](https://doi.org/10.1109/JPROC.2018.2867029).
- [3] ISLAM M T, TAHA A E M, and AKL S. A survey of access management techniques in machine type communications[J]. *IEEE Communications Magazine*, 2014, 52(4): 74–81. doi: [10.1109/MCOM.2014.6807949](https://doi.org/10.1109/MCOM.2014.6807949).
- [4] DAI Linglong, WANG Bichai, DING Zhiguo, *et al.* A survey of non-orthogonal multiple access for 5G[J]. *IEEE Communications Surveys & Tutorials*, 2018, 20(3): 2294–2323. doi: [10.1109/COMST.2018.2835558](https://doi.org/10.1109/COMST.2018.2835558).
- [5] WONG V W S, SCHÖBER R, NG D W K, *et al.* Key Technologies for 5G Wireless Systems Physical Layer Caching with Limited Backhaul in 5G Systems[M]. Cambridge: Cambridge University Press, 2017: 236–270. doi: [10.1017/9781316771655](https://doi.org/10.1017/9781316771655).
- [6] LIEN Shaoyu, HUNG Shaochou, DENG D J, *et al.* Efficient ultra-reliable and low latency communications and massive machine-type communications in 5G new radio[C]. Proceedings of IEEE Global Communications Conference, Singapore, 2017: 1–7. doi: [10.1109/GLOCOM.2017.8254211](https://doi.org/10.1109/GLOCOM.2017.8254211).
- [7] PLODER O, PALAORO N, ETZLINGER B, *et al.* A cross-layer approach for ultra-low-latency machine type communication[C]. Proceedings of IEEE International Conference on Communications, Paris, France, 2017: 1–6. doi: [10.1109/ICC.2017.7997462](https://doi.org/10.1109/ICC.2017.7997462).
- [8] SUN Xiaofang, YAN Shihao, YAN Nan, *et al.* Short-packet downlink transmission with non-orthogonal multiple access[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(7): 4550–4564. doi: [10.1109/TWC.2018.2827368](https://doi.org/10.1109/TWC.2018.2827368).

- [9] LI Anxin, CHEN Xiaohang, and JIANG Huiling. Contention based uplink transmission with NOMA for latency reduction[C]. Proceedings of the 2017 IEEE 85th Vehicular Technology Conference, Sydney, Australia, 2017: 1–5. doi: [10.1109/VTCSpring.2017.8108486](https://doi.org/10.1109/VTCSpring.2017.8108486).
- [10] 3GPP. Study on scenarios and requirements for next generation access technologies[R]. TR 38.913, 2017.
- [11] 3GPP. Further advancements for E-UTRA physical layer aspects[R]. TR 36.814 v9.0.0, 2010.
- [12] AIJAZ A, TSHANGINI M, NAKHAI M R, *et al.* Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical QoS guarantees[J]. *IEEE Transactions on Communications*, 2014, 62(7): 2353–2365. doi: [10.1109/TCOMM.2014.2328338](https://doi.org/10.1109/TCOMM.2014.2328338).
- [13] ABDELSADEK M Y, GADALLAH Y, and AHMED M H. An LTE-based optimal resource allocation scheme for delay-sensitive M2M deployments coexistent with H2H users[C]. Proceedings of IEEE Conference on Computer Communications Workshops, Atlanta, USA, 2017: 139–144. doi: [10.1109/INFCOMW.2017.8116366](https://doi.org/10.1109/INFCOMW.2017.8116366).
- [14] SHE Changyang, YANG Chenyang, and QUEK T Q S. Radio resource management for ultra-reliable and low-latency communications[J]. *IEEE Communications Magazine*, 2017, 55(6): 72–78. doi: [10.1109/MCOM.2017.1601092](https://doi.org/10.1109/MCOM.2017.1601092).
- [15] AIJAZ A. Towards 5G-enabled tactile internet: radio resource allocation for haptic communications[C]. Proceedings of IEEE Wireless Communications and Networking Conference, Doha, Qatar, 2016: 1–6. doi: [10.1109/WCNC.2016.7564661](https://doi.org/10.1109/WCNC.2016.7564661).
- [16] SHARIATMADARI H, DUAN Ruifeng, IRAJI S, *et al.* Resource allocations for ultra-reliable low-latency communications[J]. *International Journal of Wireless Information Networks*, 2017, 24(3): 317–327. doi: [10.1007/s10776-017-0360-5](https://doi.org/10.1007/s10776-017-0360-5).
- [17] DI Boya, SONG Lingyang, LI Yonghui, *et al.* Non-orthogonal multiple access for high-reliable and low-latency V2X communications in 5G systems[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 25(10): 2383–2397. doi: [10.1109/JSAC.2017.2726018](https://doi.org/10.1109/JSAC.2017.2726018).
- 谢显中: 男, 1966年生, 博士, 教授, 博士生导师, 研究方向为移动通信网络、认知无线电技术等。
- 黎佳: 女, 1993年生, 硕士生, 研究方向为MTC与高可靠低延迟通信。
- 黄倩: 女, 1990年生, 博士生, 研究方向为车联网与高可靠低延迟通信。
- 陈杰: 男, 1993年生, 硕士生, 研究方向为车联网与网络编码。