

大数据的无线传输。以数据为导向的 性能极限和它们的 应用

杨红川，*IEEE* 高级会员和Mohamed-Slim Alouini，研究员。

IEEE

摘要

大数据和物联网（IoT）应用的日益普及给无线通信界带来了新的挑战。无线传输系统应更有效地支持来自不同类型信息源的大量数据流量。在这篇文章中，我们为无线传输策略的设计和优化引入了一种新的面向数据的方法。具体来说，我们为单个数据传输会话定义了新的性能指标，并应用它们来比较两种流行的信道适应性传输策略。我们对这些传输策略提出了几个有趣的、有点反直觉的看法，这在传统方法中是不可能的。我们还提出了几个有趣的未来研究方向，值得用面向数据的方法进行研究。

索引词

大数据、物联网、无线通信、衰减信道、自适应传输、传输时间、熵值和吞吐量。

I. 简介

我们正处于一个大数据时代。数据的产生和收集速度在不断加快。这些数据的及时处理、传递和分析将带来巨大的社会和经济效益。

这项工作得到了NSERC发现资助的部分支持。

H.-C.Yang在加拿大维多利亚大学电子和计算机工程系，维多利亚，BC V8W 2Y2，（电子邮件：hy@uvic.ca）。

M.-S.Alouini在阿卜杜拉国王科技大学（KAUST）计算机、电气和数学科学与工程部（CEMSE）工作，地址是Thuwal 23955, Saudi Arabia（电子邮件：slim.alouini@kaust.edu.sa）。

ar
Xi
v:
18
05
.0
99
23
v2
[e
es
s.
SP
]
20
19
年
4
月
16
日

益[1], [2]。随着无线通信系统的密集持续部署,大多数大数据将通过空中传输。事实上,智能移动设备对大数据的产生贡献巨大。不断增长的物联网(IoT)设备成为无线传输的另一个大数据来源。对大数据传输的支持给无线系统设计带来了一些技术挑战,包括提高无线接入网(RAN)的频谱效率,提供前导/回程链路的容量,以及改进网络结构以实现流量的可扩展性。为了有效地支持各种大数据和物联网应用,未来的无线系统需要针对来自不同来源的大量数据优化其传输策略。

在过去的二十年中,数字无线传输技术有了很大的发展。各种先进的传输技术,包括多天线(MIMO/massive MIMO)传输[3], [4], 信道自适应传输[5], 合作中继传输[6], [7], 认知无线电传输[8], 以及极端带宽传输(如毫米波,太赫兹和光学无线传输),都被开发和部署以满足对高数据率无线服务日益增长的需求。这些传输技术的设计目标通常是提高或接近无线信道的容量极限,通常以误码容量和断码容量为特征。遍历容量规定了衰减信道上平均传输速率的上限,而中断容量对应于特定中断概率约束下信道的最大瞬时传输速率。其原理是,提高信道质量必然会改善单个传输会话的平均服务质量。到目前为止,这种*面向信道*的方法相当有效,并成功地促进了高质量无线服务的提供。

目前的无线系统通常对无线信道通道上的所有传输会话应用相同的传输策略。随着先进传输技术的应用,信道的属性,如平均数据率和平均错误率,将得到改善,这通常会转化为单个会话所经历的更好的平均服务质量。这种面向信道的设计对于持续时间较长的传输会话,如电话和视频流,效果非常好。同时,面向信道的设计忽略了单个传输会话的具体情况,如流量特性和当时的网络/信道条件。当传输会话很短时,单个会话所经历的服务质量在平均水平上有很大的差异。特别是。

最近有研究表明，在衰减信道上自适应传输的固定数据量的传输时间在其平均值附近有很大的变化[9]。随着物联网设备和大数据应用的日益普及，未来的无线系统需要支持越来越多的短距离传输会话，例如由传感器节点发起。

为了进一步提高无线传输系统的效率，尤其是物联网和大数据应用，我们需要从新的角度研究无线传输技术。在这篇文章中，我们主张从单个传输会话的角度来研究。直观地说，我们期望，如果根据流量特性和运行环境对每个传输会话的传输策略进行优化，那么无线传输的性能/效率可以进一步提高。在这种直觉的激励下，我们提出了一种新的*面向数据*的无线传输系统设计方法。具体来说，当有一定量的数据可供传输时，我们将以最优方式决定传输策略。例如，功率适应是否应该与速率适应一起应用？是否应该启动合作中继？应该采用何种多天线传输结构？传输策略将根据流量特性和信道/网络条件为每个数据传输会话进行调整。以数据为导向的方法的理由是，优化单个会话的传输策略将直接改善它们的服务质量，并将反过来提高整个系统的传输效率。我们相信所提出的面向数据的方法将促进大数据和物联网应用的高效传输解决方案的设计。

对于新的面向数据的方法，有许多挑战需要解决。我们首先需要定义合适的指标来量化单个数据传输会话所经历的服务质量。我们还需要从数据传输的角度建立性能限制，并将其作为优化传输策略的准则。在这篇文章中，我们提出了一些面向数据方法的初步调查。特别是，我们为单个数据传输会话引入了两个新的面向数据的性能限制。作为这些性能限制的初步应用，我们比较了两种流行的信道自适应传输策略在衰减信道上的性能，当信道状态信息在发射器处可用时。最后，我们讨论了面向数据的方法的几个有前途的应用和未来研究方向。

II. 面向数据的性能限制

误码率和断码率是众所周知的在衰减信道上进行无线传输的性能极限。频谱容量适用于传输将经历所有可能的衰减状态的情况。它描述了信道在快速衰落环境或极长的传输时间内所能支持的最大传输速率。另一方面，中断容量适用于慢速衰减环境，并规定了在特定中断概率要求下信道可以支持的最大传输速率。一般来说，中断容量规定了瞬时容量限制，即在信道相干时间内，信道的实现是高度相关的，而遍历容量规定了在很长的时间内的平均速率限制，例如比相干时间大几个数量级。这些*面向信道*的性能限制不能完全描述单个数据传输会话所经历的服务质量，特别是对于大数据和物联网应用。

大数据的无线传输往往涉及多个信道的相干时间。例如，考虑通过IEEE 802.11ac WiFi进行AR/VR视频的室内传输。AR/VR视频的典型文件大小约为4Gbits，而802.11ac的峰值下载速度可以达到2.5Gbps。因此，视频传输可以在平均1.6秒内完成。WiFi典型工作环境的信道一致性时间约为200毫秒。因此，传输将持续约八个相干时间段。作为另一个例子，考虑通过LTE链路进行高质量图像的户外传输。图像的文件大小在压缩后可以达到几百Kbits，LTE链路的传输速度可以达到Mbps。传输将持续大约几百毫秒，这需要几个相干时间段，而户外环境的典型相干时间值为几十毫秒。

频谱容量只能表征一个特定的无线传输会话在平均意义上经历的服务质量。数据传输会话的实际传输服务在很大程度上取决于当时的信道实现。特定会话的有效传输率在平均传输率附近会有很大的变化。为了更有效地描述传输服务的质量，我们提出以下问题。给定要传输的一定数量的数据，在固定的时间内成功传输的机会是多少？鉴于可用的时间-频谱资源，什么是

是指可以在信道上可靠地传输的最大数据量？断电容量的特点是瞬时速率限制，只适用于持续不到一个信道相干时间的传输会话。断电容量不能泛化到跨越多个信道相干时间的传输。为了有效解决上述设计问题，我们需要合适的*面向数据*的新性能指标。在下文中，我们提出了两个新的性能限制。

A. 最小传输时间

许多大数据和物联网应用的基本服务要求是将一定量的数据及时传输到目的地。因此，我们定义了一个面向数据的指标--

最小传输时间（MTT），作为在无线信道上传输一定量的数据所需的最小时间。让 H 表示要传输的数据量。MTT将是 H 的函数，用 $T_{\min}(H)$

表示。对于一个给定的 H 值，MTT将改变

伴随着信道带宽、信道实现和采用的传输策略。当

H 相对较小，数据传输在一个通道相干时间内完成，MTT

T_{\min}

(H) 取决于瞬时信道的实现。通过最优速率适应（ORA）[10]，信道相干时间内的最大传输速率等于 $B \cdot \log_2$

$(1 + \gamma)$ ，其中 B 是信道带宽， γ 是瞬时接收信噪比。那么MTT可以是

计算为 $H/B \log_2(1 + \gamma)$ ，它将随接收信噪比 γ 的变化而变化。另一方面，当 H 非常大且数据传输涉及许多相干时间时，MTT可以是

使用信道的遍历容量计算，由 $C = \int_0^\infty B \log(1 + \gamma) p(\gamma) d\gamma$ 给出。因为 $MTT = H/C$ ，这是一个恒定的值。

为了解决前面的设计问题，我们将延迟中断率（DOR）定义为一定数量的数据的MTT大于阈值时间的可能性。特别是，DOR在数学上由 $DOR = \Pr[T_{\min}(H) > T_{th}]$ 给出，其中 T_{th} 表示阈值

持续时间。在信息理论意义上， H 代表了包含的信息量。

在数据中。 T_{th}

可以与要传输的数据的延迟要求有关。因此，DOR可作为单个数据传输会话所经历的服务质量的一个统计措施。例如，在一个信道相干性内的数据传输的DOR

与ORA的时间可以计算为

$$DOR^{ora} = \Pr \left(\gamma < \exp\left(\frac{H \ln(2)}{B T_{th}}\right) - 1 \right), \quad (1)$$

其中规定了无功率适应的传输时间的性能下限。当数据传输持续一个以上的信道相干时间时，如大数据传输的情况，DOR分析变得更加具有挑战性。假设在块状消逝上的ORA。

其中，接收到的信噪比在每个相干时间 T_c 上保持不变，并变化为一个后的独立值，MTT小于 $L - T_c$ 如果 $\Pr \left[\sum_{l=1}^L T_c B \log(1 + \gamma_l) > H \right]$ ，其中 γ_l 为

the received SNR over the l th T_c . As such, the DOR for the case of $T_{th} = LT_c$ can be calculated as

$$\text{DOR}^{\text{ora}} = \Pr \left[\sum_{l=1}^L T_c B \log_2(1 + \gamma_l) > H \right]. \quad (2)$$

为了准确评估上述概率，我们需要 L 个独立随机变量之和的统计分布 $T_c B \log_2(1 + \gamma_l)$ ，这可能是用Fox

H 函数[11]可以解决的。一般情况下的DOR分析将是一个有趣的研究问题以供进一步调查。

B. 最大熵值吞吐量

无线通信系统通过分配一定的频谱-时间资源来满足大数据和物联网应用的服务要求。对某一频谱-时间资源块上可成功传输的数据量的表征，将有助于资源分配算法的设计。因此，我们将最大熵吞吐量（MET）定义为在一定的时间长度和信道带宽上可以传输的最大信息量。在数学上，我们用 $H_{\max}(T, B)$ 表示MET。它是时间长度 T 和信道带宽 B 的一个函数。一个任意的时间长度，其值从小于一个相干时间 T_c 到许多 T_c 's。对于给定的 T 和 B 值，MET将取决于信道的实现和采用的传输策略。例如，当 T 比 T_c 大一个数量级时，MET可以

可以计算为 $H_{\max}(T, B) = TC$ ，其中 C 是信道的遍历容量。另一方面，如果 T 小于 T_c ，并且应用ORA，那么MET可以用瞬时信道容量计算为 $H_{\max}(T, B) = TB \log_2(1 + \gamma)$ 。对MET的分析是在 T 跨越多个 T_c 的情况下， c ，将涉及更多。

由于MET一般随信道实现而变化，我们将信息中断率（IOR）定义为在一定时间内MET小于阈值熵的概率，用 H_{th} 表示。在数学上，IOR由 $\Pr[H_{\max}(T, B) < H_{th}]$ 给出。IOR分析

将有助于设计有效的资源分配算法。对于有严格时延要求的数据流量，最好能分配足够的时间-频谱资源，使数据传输能在时延限制内以高概率成功完成。显然，IOR分析需要 $H_{\max}(T, B)$ 的统计数据，这取决于信道的情况统计数据和采用的传输策略。例如，当 T 非常大时，相对于 T_c ，如果信道遍历容量 C 大于 H_{th}/T ，IOR将等于0。同时，当 T 小于 T_c ，假设系统采用ORA，IOR可以计算为

$$\text{IOR}^{\text{ora}} = \Pr[\gamma < \exp(\frac{(\ln(2)H_{\text{th}})}{T}) - 1]。 \quad (3)$$

对于 T 涉及多个 T_c 的情况，IOR将等于MET小于 H_{th} 的概率，其评估将需要 N 个独立随机变量之和的分布。对IOR的进一步研究将是未来一个有趣的话题。

研究。

III. 与CSIT的传输策略比较

作为上一节介绍的面向数据的性能限制的应用，我们现在比较了点对点无线信道上的ORA和最优功率和速率适应（OPRA）策略。当发射器上有完整的信道状态信息时（CSIT），采用ORA的无线传输可以达到消隐信道的遍历容量。也有学者认为，OPRA传输可以进一步提高有水填充功率分配的消逝无线信道的容量[10]。特别是，在低信噪比区域，所产生的OPRA容量大大高于遍历容量[12]。从一个特定的数据传输会话的角度来看，OPRA传输仍然会优于ORA传输吗？我们可以应用上节介绍的面向数据的性能指标，即DOR和IOR，来回答这个问题。

在慢速衰减环境下的DOR和IRR分别由公式（1）和公式（3）给出。在OPRA下，当接收信噪比大于阈值信噪比 γ_T ，瞬时信道容量为 $B \log_2(\gamma/\gamma_T)$ ，否则为零。阈值信噪比 γ_T 是确定，以满足平均发射功率的约束，并采用最佳充水功率分配政策。因此，在慢速衰减环境下，采用OPRA的DOR计算为

$$\text{DOR}^{\text{opra}} = \Pr[\gamma < \exp(\frac{H \ln(2)}{\gamma_T})], \quad (4)$$

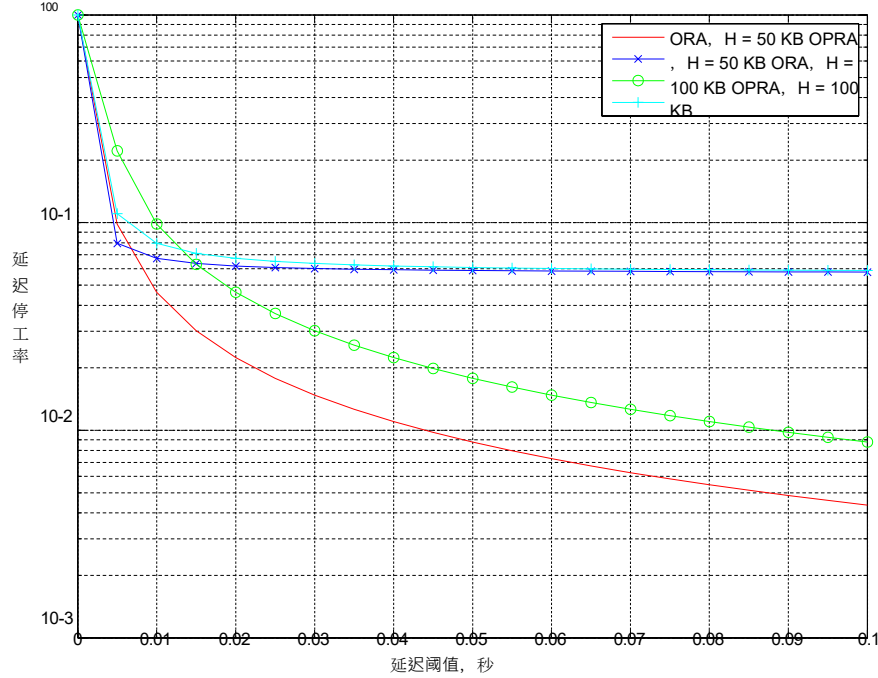


图1.ORA和OPRA在慢速Rayleigh衰减信道上传输的延迟中断率 ($B=20\text{MHz}$, $\gamma=6\text{dB}$)。

而OPRA在慢速衰减信道上传输的IOR被确定为

$$\text{IOR}^{\text{opra}} = \Pr[\gamma < \exp(\frac{Hth \ln(2)}{\gamma T})]。 \quad (5)$$

图1比较了ORA和OPRA传输策略在慢速Rayleigh衰减信道上的DOR性能。特别是，我们绘制了两种策略的DOR与不同数据量 H 的延迟阈值 T_{th} 的函数。我们可以看到，对于两种 H 值的选择，都是ORA和OPRA的DOR性能之间的混合行为。具体来说，当延迟阈值很小，OPRA导致的DOR比ORA小。当阈值持续时间变大时，使用ORA传输的DOR会有所改善，变得比使用OPRA的DOR小很多。事实上，当延迟阈值变得非常大时，OPRA的DOR会收敛到一个固定值，这等于OPRA不传输的概率。图2说明了平均接收SNR对DOR性能的影响。我们可以看到，当平均信噪比很小的时候，ORA总是比OPRA取得更小的DOR，这就保证了传输的概率更高。当平均信噪比增加时，OPRA的DOR性能有所提高，但当延迟阈值较大时，仍然比OPRA差。请注意，从传统的遍历容量的角度来看，OPRA的性能大大超过了ORA。

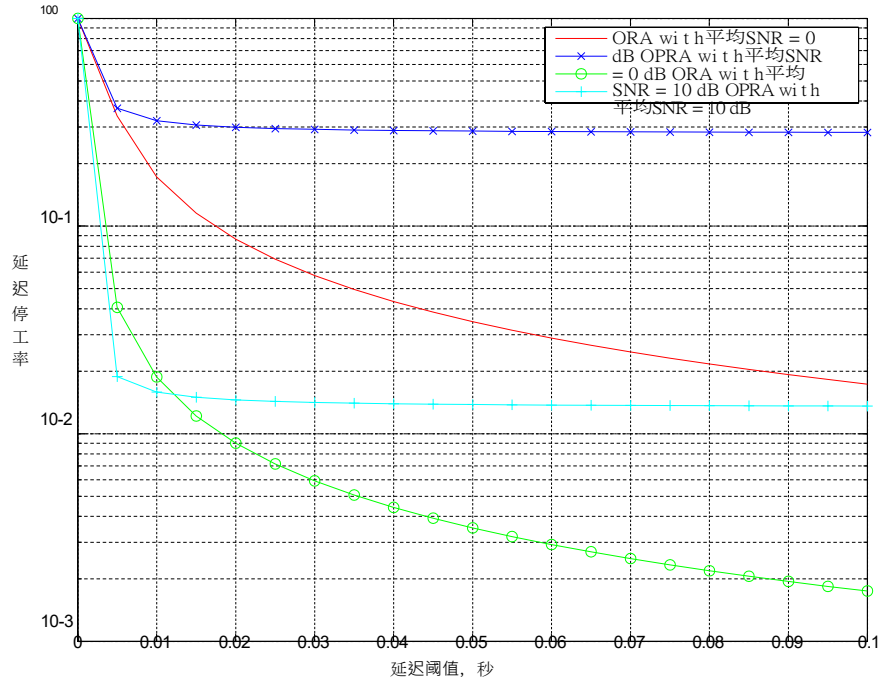


图2.平均信噪比对慢速雷利消逝信道上ORA和OPRA传输的延迟中断率的影响 ($B=20\text{MHz}$, $H=50\text{KB}$)。

低信噪比制度。然而，我们从DOR的比较中观察到，从单个传输会话的角度来看，OPRA并不总是更好的策略。当延迟要求非常严格或信道质量很好时，OPRA比ORA更受欢迎。

我们现在比较ORA和OPRA在慢速Rayleigh衰减信道上的传输策略的IOR性能。在图3中，我们绘制了两种策略的IOR与熵的函数关系。

阈值 H_{th}

，不同的时间长度 T 。我们再次观察到ORA和OPRA的IOR性能之间的混合行为。具体来说，当熵阈值较小时，ORA

导致的IOR值比OPRA小。当熵阈值变大时，采用ORA传输的IOR就会增加，并迅速变得比采用OPRA的IOR大。事实上，OPRA的IOR从一个固定值开始稳步上升，该值等于不传输的概率，对于两个时间长度值来说都是如此。图4说明了平均接收信噪比对IOR性能的影响。我们可以看到，当平均信噪比很小的时候，在更大的熵阈值范围内，ORA在IOR方面优于OPRA。当平均信噪比增加时，OPRA的IOR性能明显改善，并优于ORA的性能。

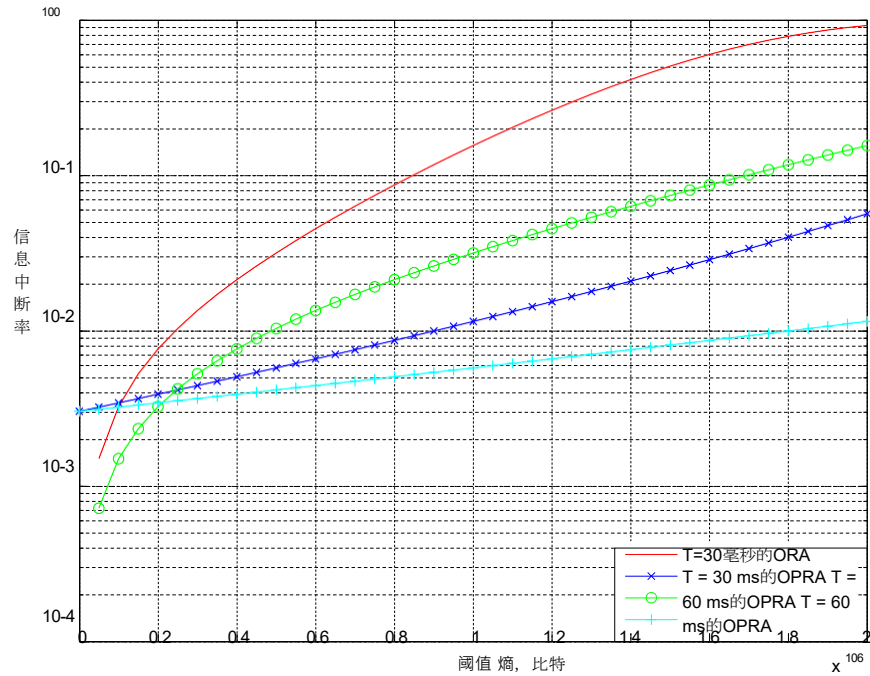


图3.ORA和OPRA在慢速Rayleigh衰减信道上传输的信息中断率 ($B = 20$ MHz, $\gamma = 6$ 分贝)。

除非 H_{th}

极小。从这个比较中我们可以得出结论，从单个传输会话的角度来看，OPRA传输策略应该在高信噪比上使用

场景和/或需要传输大量数据时。

IV. 进一步的应用

新的面向数据的性能极限表征了在消逝的无线信道上单个数据传输会话的性能。特别是，MTT规定了在衰减信道上传输一定数量的数据时可能的最小传输延迟，而MET则表示在一个时间-

频谱资源块上可以传输的最大信息量。鉴于无线衰落信道的时变性，这些性能极限是在统计学意义上描述的，分别是DOR和IOR。通过规定单个传输会话的最佳性能，这些性能极限将在无线传输策略的设计和优化中找到许多重要的应用。

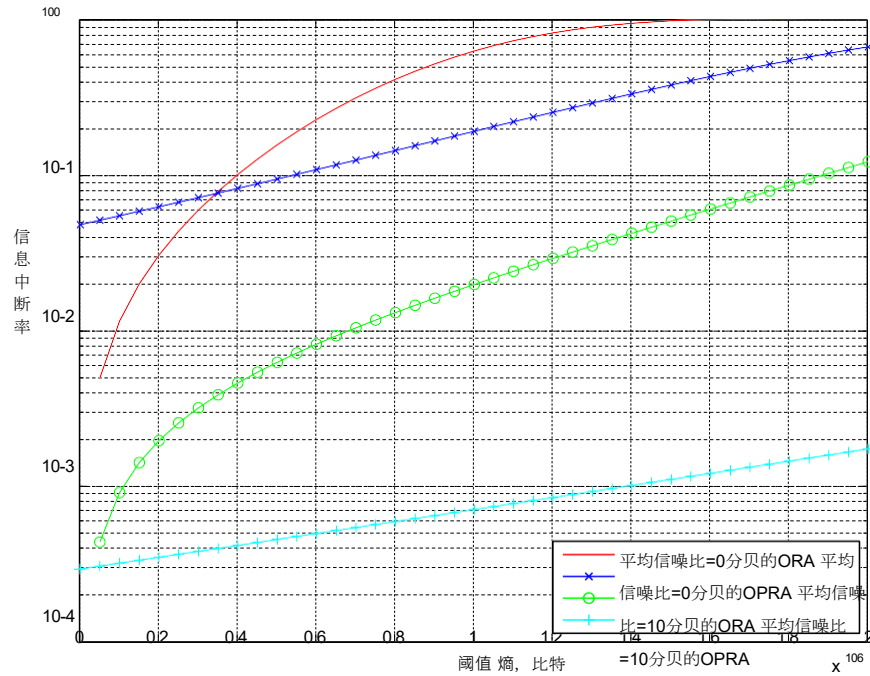


图4.平均信噪比对慢速雷利消散信道上ORA和OPRA传输的信息中断率的影响 ($B = 20$ MHz, $T = 30$ ms)。

在上一节中，我们使用面向数据的指标，比较了慢速衰减情况下CSIT的两种信道自适应传输策略。其中的设计见解可以很容易地应用于物联网流量的传输方案优化，这些流量通常是短暂和零星的。另一方面，大数据应用往往会产生大量的数据流量，其传输可能持续多个信道相干时间。针对大数据流量的数据导向分析将是一个具有挑战性但有价值的未来研究课题。最近有报道称，对大数据的传输时间进行了初步调查，并在消退信道上进行了离散速率适应[9]。

面向数据的方法也可以适用于设计和优化具有有限CSIT的实际传输策略。自适应调制和编码 (AMC) 和自动重复请求 (ARQ) 是两种流行的传输策略，它们探索来自接收器的有限反馈。AMC根据一定的可靠性要求调整传输速率，而ARQ则通过重传提高可靠性。文献[13]已经研究了AMC和ARQ的联合设计。通过提出的面向数据的方法，我们可以研究这些

从一个全新的角度来研究两种传输策略及其联合设计。这样的研究将产生新的设计见解，并导致在有限的CSIT情况下产生新的传输策略。MTT分析表征了实际数据传输的传输时间的下限。对于点对点链接，传输时间与传输系统的服务速率成反比。无线传输的排队延迟性能可以用传输时间的一阶和二阶统计来分析[14]。我们以数据为导向的描述可以用于开发点对点链路的排队性能的上界。同时，传输时间与每个传输会话的信道占用率直接相关。传输的统计表征可用于优化随机访问协议。

大数据和物联网应用的日益普及将创造出前所未有的流量和不同的服务要求。无线系统需要应用有效的资源分配算法，以有效地适应这种新的需求。MET表征将为资源分配算法的设计和优化提供宝贵的指导。例如，3GPP为物联网规定采用了预定上行链路的方法，这涉及到调度请求的随机接入阶段[15]。只有在这个阶段成功的终端才会被分配到资源块。因此，为每个终端分配足够的资源块是非常关键的，这样他们的传输才能高概率地完成。通过面向数据的方法，我们可以提高这种资源分配算法的性能。

V. 结论性意见

在这篇文章中，我们提出了一种新的面向数据的无线传输系统设计和分析方法。我们的目标是根据流量特性和工作条件，对单个数据传输会话进行传输策略设计和优化。特别是，我们引入了两个面向数据的性能限制来描述任意的无线数据传输。作为其最初的应用，我们比较了CSIT场景下著名的信道自适应传输策略，即ORA和OPRA。我们观察到，虽然从遍历容量的角度来看，OPRA总是优于ORA，但从单个数据传输会话的角度来看，OPRA并非总是首选的传输策略。在慢速衰落信道上，当平均功率为0.5W时，OPRA可以有更好的机会将数据传送到目的地。

信道质量很差。因此，面向数据的方法可以为衰减信道上的无线传输带来有趣的新见解。

这篇文章是对面向数据的无线传输策略设计方法的初步介绍。有许多重要的方面需要解决。发射器处有限的和没有CSI的情况是有实际意义的。面向数据的大数据传输的特征需要进一步调查。最后，这些表征可以随时应用于无线系统的排队分析和资源分配。我们期望，面向数据的设计将大大促进大数据和物联网应用的高效无线传输策略的设计。

参考文献

- [1] T.Hu, E. Bigelow, J. Luo, and H. Kautz, "Tales of two cities:使用社交媒体来了解独特的大都市地区的特异性生活方式, " *IEEE Transactions on Big Data*, vol. 3, no. 1, pp. 55-66, Mar. 2017.
- [2] F.Xia, W. Wang, T. M. Bekele, and H. Liu, "大型学术数据。A survey," *IEEE Transactions on Big Data*, vol. 3, no. 1, pp.18-35, Mar. 2017.
- [3] G. J. Foschini和M. Gans, "关于在衰减环境中使用多天线时的无线通信限制", *Wireless Pers. Commun.*, pp. 311-335, Mar. 1998.
- [4] E.Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, 第52卷, 第2期, 第186-195页, 2014年2月。
- [5] S.-G.Chua and A. J. Goldsmith, "Adaptive coded modulation for fading channels," *IEEE Trans.通信。*, vol. 46, no.5, pp.595-602, 1998年5月。
- [6] A.Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description, " *IEEE Trans.Wireless Commun.*, 第51卷, 第11期, 第1927-1938页, 2003年11月。
- [7] J.N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks:高效协议和停电行为", *IEEE Trans.Inf.Theory*, 第50卷, 第11期, 第3062-3080页, 2004年11月。
- [8] J.Mitola和G. Q. Maguire, "认知无线电:使软件无线电更个人化", *IEEE Trans.无线通信。*, vol. 6, no.4, pp. 13-18, Aug. 1999.
- [9] W.-J. Wang, H.-C.Yang, and M.-S.Alouini, "大数据的无线传输。渠道上的传输时间分析", *IEEE Trans.Wireless Commun.*, pp.1-11, 2018.
- [10] A.Goldsmith, *Wireless Communications*. 剑桥大学出版社, 2005年。
- [11] F.Yilmaz和M. S. Alouini, "多载波系统的中断能力", 在2010年第17届国际电信会议上, , 2010年4月, 第260-265页。
- [12] Z.Rezki and M. S. Alouini, "On the capacity of Nakagami-m fading channels with full channel state information at low SNR," *IEEE Wireless Communications Letters*, Vol. 1, no.3, pp. 253-256, June 2012.
- [13] Sassioui, M. Jabi, L. Szczecinski, L. B. Le, M. Benjillali, and B. Pelletier, "HARQ和AMC。朋友还是敌人?" *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 635-650, Feb 2017.
- [14] M.Usman, H. C. Yang, and M. S. Alouini, "Extended delivery time analysis for cognitive packet transmission with application to secondary queuing analysis, " *IEEE Trans.无线通信。*, 第14卷, 第10期, 第5300-5312页, 2015年10月。

- [15] Y.P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3GPP narrowband Internet of things, " *IEEE Commun.Mag.*, vol. 55, no.3, pp. 117-123, March 2017.