# 基于空中计算的可解释性联邦学习中功率控制算法

## 引言

随着移动网络快速发展，智能终端规模也随之急剧增长，传统的云计算模式无法满足移动设备的低延时和隐私需求，因此移动边缘计算被广泛研究并应用。移动边缘计算模式通过利用边缘节点与智能终端的计算能力，在靠近数据生成的位置来处理、分析、存储数据，从而提高数据的处理速度，并带来数据隐私保护和安全性。联邦学习是一种新的分布式机器学习架构，各边缘节点或智能终端先利用自身的数据训练本地模型，再通过交互模型中间参数进行模型联合训练。这种分布式方法的优势在于它更优秀的模型、更小的能耗与更低的延迟，并且还能保护数据隐私。

联邦学习的主要挑战之一是频繁的模型参数传输汇总，海量节点的接入将会导致通信频域分配减少和带宽降低，从而将影响联邦学习训练速度等性能。空中计算技术是解决这一问题有效手段，其核心思想是利用无线多址接入信道的波形叠加特性以及多用户的并发传输以实现高速数据空中聚合。它可以有效利用频谱资源，接入时延不再会随着网络规模的增加而线性增加，并且进一步地保护了用户的隐私安全。考虑到部分终端设备可能距离边缘服务器较远，信道条件较差，进而影响到空中计算性能。因此，本文将对终端设备进行分簇聚类，并增设中继器节点以保证较好的信号传输性能。

本文提出的应用空中计算的多层次联邦学习架构，即终端-中继器-边缘服务器的三层架构，在模型汇聚时利用空中计算技术，并采用中继策略放大与转发信号，使所有设备共享全频域资源，增大实际带宽，减小传输时延，降低信号失真率。在此基础上设计了基于可解释性的模型聚类算法，可以依据可解释性算法对一组聚类样本集进行解释，利用该解释分析模型推理的内在原理并得到该模型的可解释性矩阵，将该矩阵作为聚类的参考指标实现模型分簇，有效地对初始模型进行分类。同时，详细分析出空中计算在通信过程中特点，提出在噪声影响下满足功率限制下信号失真率优化模型，并通过调节发射器的发射功率与接收器接收因子最小化信号失真率。最后，通过理论验证了出两跳联邦学习的失真率小于一跳联邦学习的情况，同时仿真结果中，对比了其他主流一跳联邦学习算法，本方法在准确率与收敛速度等方面表现更为出色。

本章的剩余部分组织如下。第二节介绍了多层联邦学习架构与在该场景下信号失真率优化问题。第三节介绍了基于可解释性的聚类算法与数学模型的迭代求解算法，并且理论验证两跳联邦学习的失真率小于一跳联邦学习的情况。第四节分析了仿真结果。最后总结了本章的结论。

## 多层联邦学习模型

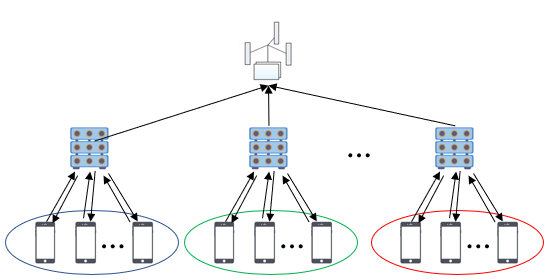
由于部分终端到边缘服务器的信道条件差，会明显影响信号传输的失真率，因此本课题考虑通过增加传输跳数，设立中继节点来中继放大并转发，有效减小信号的传输失真率。本课题提出的一种应用空中计算的多层次联邦学习架构，即终端-中继器-边缘服务器的三层架构。终端-中继器层利用空中计算完成信号传输，有效地共享频域资源，减低传输时延，中继器再放大并转发信号到边缘服务器完成联邦学习，并通过调节功率有效降低传输信号的失真率。

图4-1 应用空中计算的可解释联邦学习架构

如图4-1所示，本研究的边端协同场景中共有N个终端，R个中继器与一个边缘服务器。其中终端n∈N将依据可解释性的方法分簇，各簇的终端设备将选择对应的中继器完成信号的放大转发，再由中继器发送信号到边缘服务器，从而实现两跳信号传输。在第一跳中，空中计算组中有个终端，它们将同时发送本地训练模型参数信号，在同一信道下利用空中计算的方案传输到中继器，同时满足各终端的传输功率峰值约束。每个在相同空中计算组的终端将会复用正交子信道k并将信号传输给中继器r。中继器将接收空中计算组中不同终端的叠加信号，并将叠加信号放大转发给基站。

第一阶段中，由于中继与空中计算组中的终端复用信道，其中的个终端的信号将会被另一个中继干扰，因此中继r中收到空中计算组的叠加信号为：

其中表示从终端n到中继r的信道增益，表示从中继o到中继r的信道增益。表示终端n的发射系数，是终端n发射的预处理完的信号（均值为0，方差为1），表示簇外某中继器的发射系数，是该中继器发射的信号，是中继r的加性高斯白噪声（AWGN）。

其中，终端的发射系数将满足功率约束：

第二阶段中，中继器r使用边缘服务器分配的子信道，使用放大并转发策略将接收到的信号传输到边缘服务器，因此边缘服务器收到的信号为：

其中表示从中继r到基站的信道增益，表示中继器的发射系数，是中继器发射的信号，是边缘服务器的加性高斯白噪声（AWGN）。

边缘服务器接收到信号后，会对信号进行降噪处理，实际接收信号值为：

其中表示接收因子，可以降低噪声对传输信号的影响。表示第一阶段中簇内终端的数量。

当不考虑任何干扰与噪声，在理想情况下的理论接收信号值为：

对于实际信号值与理论信号值之间的差别，本研究利用均方误差来衡量信号的失真率，均方误差表示为：

本课题将调节发射因子与接收因子，从而达到降低信号传输失真率的目的，因此优化问题为：

针对优化问题(4-7)，本课题采用一种基于迭代的求解算法，将问题分解为两个子问题，分别求出，的最优值表达式，不断通过迭代的方式，逼近最优解，使最终MSE的变化率小于阈值，输出，得到最终结果。具体求解算法将在4.3节中阐述。

## 求解算法

针对优化问题(4-7)，本课题首先将基于可解释性算法的聚类算法来对终端节点进行分簇，确定簇内节点组的节点并接入中继节点实现通信，形成终端-中继节点-边缘服务器的两跳信号传输架构，实现对中继辅助的空中计算策略的支持，使所有设备共享全频域资源，增大实际带宽，减小传输时延，降低信号失真率。之后，本课题将所构建的目标问题拆分，分解为发射因子优化和接收因子优化两个子问题，并利用迭代的方式获取最优值，实现对目标问题的求解。

### 基于可解释性的模型聚类方法

本课题提出了一种基于可解释性的模型聚类方法，终端节点的模型根据本地数据预训练，构建本地的预测模型，各模型将依据可解释性算法对一组聚类样本集进行解释，利用该解释分析模型推理的内在原理得到该模型的可解释性矩阵，将该矩阵作为指标实现模型的分簇。

算法的具体步骤如表4-1所示。由于终端节点所处的地理位置不同，收集的本地数据也将会有一点差异。各终端会基于自身的独立的本地数据集进行模型训练，从而构建出初始本地模型。各终端将使用一组相同的聚类样本集进行聚类分簇。分簇的过程中，终端将会利用可解释性算法Anchors对所有的样本进行解释，得到对应的本地解释集，如表4-2所示。各行的解释是代表预测的充分条件，即模型做出预测是依据这些特征完成的，因此无论对这一充分条件之外的特征做任何改动，预测结果都不会受到影响。为了更方便后续的聚类分簇，本文对这个本地解释集进行一定的处理，使矩阵中的值只含有0和1。如表4-3所示，当解释中含有样本i的属性j时，矩阵的第i行j列为1；否则，如果解释中不含该属性，矩阵对应的值转化为0。将所得到的解释矩阵作为聚类的指标，利用K-means聚类算法完成最终的分簇，实现基于可解释性的模型聚类算法。

表4-1 基于可解释性的模型聚类算法过程

|  |
| --- |
| Input: 各终端的独立本地数据集  Output: 聚类结果 |
| 1：终端节点的模型基于本地数据完成预训练，构建本地模型 |
| 2：各终端节点获取一组相同的聚类样本集 |
| 3：终端节点利用可解释性算法，对聚类样本集进行预测的解释，形成本地解释集 |
| 4: 终端节点将本地解释集转化为解释矩阵 |
| 5：终端节点将解释矩阵作为聚类指标，利用K-means聚类算法分簇 |
| 6：得到相关聚类结果 |

表4-2 本地解释集

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **花萼宽度** | **花萼长度** | **花瓣宽度** | **花瓣长度** |
| **样本1** | **>1.8** |  | **<2.8** | **>2.4** |
| **样本2** | **>2.1** |  |  | **>2.2** |
| **样本3** |  | **<2.5** | **<1.6** | **>2.1** |
| **样本4** |  | **<1.6** | **>1.7** |  |
| **……** | **……** | **……** | **……** | **……** |

表4-3 解释矩阵

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 行\列 | 1 | 2 | 3 | 4 |
| 1 | 1 | 0 | 1 | 1 |
| 2 | 1 | 0 | 0 | 1 |
| 3 | 0 | 1 | 1 | 1 |
| 4 | 0 | 1 | 1 | 0 |
| …… | …… | …… | …… | …… |

### 空中计算信号失真率

部分终端到边缘服务器的信道条件差，会明显影响信号传输的失真率，因此本研究将考虑通过增加传输跳数，设立中继节点来中继放大并转发，有效减小信号的传输失真率。

接下来将证明增设中继情况下两跳传输信号MSE会小于一跳传输信号的MSE。在一跳传输的情况下，基站接收到的空中计算叠加信号为：

其中表示从终端n到边缘服务器的信道增益，表示终端n的发射系数，是终端n发射的预处理完的信号（均值为0，方差为1），是中继r的加性高斯白噪声（AWGN）。

边缘服务器接收到信号后，会对信号进行降噪处理，实际接收信号值为：

其中表示接收因子，可以降低噪声对传输信号的影响。表示簇内终端的数量。

当不考虑任何干扰与噪声，在理想情况下的理论接收信号值为：

对于实际信号值与理论信号值之间的差别，一跳传输的均方误差表示为：

因此，证明两跳传输信号MSE会小于一跳传输信号的MSE，即：

研究中假定所有的噪声均为加性高斯白噪声（AWGN），因此可以假定，因此上式转化为：

由于两跳传输中增设了中继，中继与终端之间的信道条件较好，不会出现边缘服务器与部分终端之间信道过差的情况，并且很小，噪声影响不大，因此在式中的左边部分可以忽略，则

为了更清楚展示证明，本文记，通过移项得：

合并化简后得到：

记Ng中的终端到中继的信道增益平均值为，终端到边缘服务器的信道增益平均值为，终端发射因子的平均值为，上式可转化为：

即

可以观察到该式为一元二次不等式，当该不等式的△>0时有解，因此需要计算不等式的△：

化简后得到：

令。原问题转化为时，存在解，因为,并且比大好几个数量级，所以上式转化为：

又因为，，所以，

同理，。上式转化为：

因为，即

因此，当时，，原不等式存在解，即

此时的取值范围为：

因为，,所以不成立

又因为所以

综上，当时，两跳传输信号的MSE将会小于一跳的MSE，此时的取值范围为：

### 功率控制算法

针对优化问题，本课题采用一种基于迭代的求解算法，将问题分解为两个子问题，分别求出，的最优值表达式，通过迭代的方式逼近最优值，使最终MSE的变化率小于阈值，输出，得到最终结果。

#### 接收因子优化

给定，本节重新构造原问题，通过优化，得到以下子问题：

为了便于后续计算，设置变量，令，则上式可以转化为凸二次问题：

由于凸二次问题的特性，通过设置上式目标函数的一阶导为零，本方法可得到上式问题的最优解：

因此，由，可以得到一项推论：

推论：当给定时，问题4-28的最优解为

从推论中可以发现，当噪声变大时接收因子的最优值也随着变大。这是因为如果噪声变大了，接收器需要更大的接收因子来抑制噪声产生的误差，否则最终的MSE将会变大。

#### 发射因子优化

给定，本节重新构造原问题，通过优化，得到以下子问题：

其中目标函数中的常量,,被忽略。在这种情况下，可以观察到该问题是一个凸线性约束二次规划问题，由于上式问题与其对偶问题之间存在强对偶性，因此本节将利用拉格朗日对偶方法以最优地求解上式问题。设表示上式中第n个约束相关联的对偶变量，让表示约束对应的对偶变量，拉格朗日乘子法转化问题为：

令一阶偏导为0，通过应用KKT条件，得到如下结果：

推论：当给定时，问题4-32的最优解为

#### 收敛与复杂度分析

本章通过交替求解子问题(1)(2)来求解原问题，直到最终收敛。算法细节在下表4-4中总结。算法的收敛性由以下命题证明：

推论3：应用如下表4-4中的算法，问题P的函数目标值将随着迭代次数的增加而减少，直至最终收敛。

证明：本节将记作问题P中可行解的目标函数值。如表中展示的，子问题(2)的可行解也是子问题（1）的可行解(例如也是子问题(1)的可行解)。接下来本文将记和分别为问题p的第i和i+1次迭代的可行解。

如算法中第4步所示，由于给定了，将是子问题(1)的最优解，因此有

同理，由于算法中第5步由于给定了，将是子问题(2)的最优解，因此有

因此可以得出

这说明问题P的目标值在迭代过程中总是递减的，因此算法具有收敛性，推论证明完成。

表4-4 迭代算法过程

|  |
| --- |
| Input: 初始发射因子，收敛准确率  Output: 最优发射因子，最优接收因子 |
| 1：Initialization：令 |
| 2：Repeat: |
| 3： 令𝑖 = 𝑖+1 |
| 4： 给定，求解子问题(1)更新 |
| 5: 给定，求解子问题(2)更新 |
| 6：Until： |

接下来，本章将探究算法每次迭代的时间复杂度。具体来说，算法的第4步的时间复杂度为O(1),第5步的时间复杂度为O()。因此，本章所提出的算法时间复杂度由第5步主导，每次迭代的总时间复杂度为O()，是一种低复杂度和高效的算法。

## 仿真结果分析

本章在表4-5中列出了仿真中的一些实验设置。本文关注边端协同场景的服务区域为[0,1000]m×[0,1000]m的正方形区域。其中边缘服务器位于坐标，中继随机分布在边缘服务器周围，与边缘服务器的距离为100~300m。终端主要随机分布在服务区的边缘位置，距离边缘服务器距离为200~500m。仿真中设置了参考距离10m的路径损失为-60dB，并将噪声功率均设置为-80dBm。假设每个终端都具有相同的功率峰值预算20dBm，中继的平均发射功率为10dBm。

表4-5 参数表

|  |  |
| --- | --- |
| 仿真参数 | 值[单位] |
| 终端到边缘服务器距离 | 200~500[m] |
| 终端到中继距离 | 50~200[m] |
| 中继到边缘服务器距离 | 100~300[m] |
| 参考距离10m的路径损耗 | -60[dB] |
| 噪声功率, , | -80[dBm] |
| 终端最大发射功率 | 20[dBm] |
| 中继平均发射功率 | 10[dBm] |
| 迭代算法的准确率 | 0.001 |

本节将所提出的方法与下列基准方法进行比较：

* 经典联邦学习(FedAvg)：来自文献[9]和文献[10]，是 FedAvg 算法。本地终端负责训练本地数据以获得本地模型，边缘服务器通过对本地模型进行平均加权获得全局模型。
* 应用空中计算的一跳传输联邦学习(FedAvg\_Aircomp)：来自参考文献[63]，基于FedAvg 算法，在终端上传本地模型到边缘服务器时，应用空中计算计算，使信号在传输过程中就完成参数聚合。
* 以距离为分簇指标的两跳联邦学习(FedAvg\_Dis)：与本方法均为两跳信号传输，设置中继节点提高信号传输质量，以终端到中继的距离为指标完成分簇，各簇的终端设备将选择对应的中继器完成信号的放大转发，再由中继器发送信号到边缘服务器，从而实现两跳信号传输。

### 收敛速度

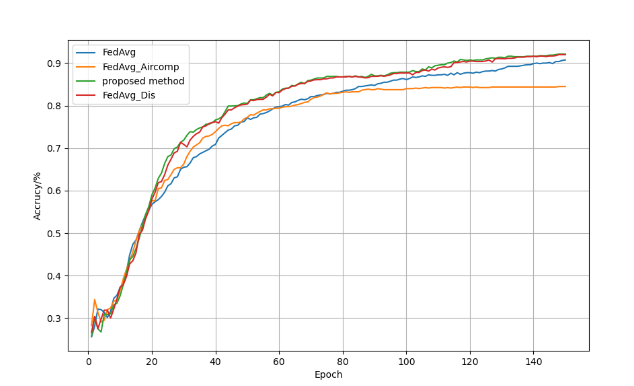
图4-2展示本方法与其他基准方案的收敛速度对比。如图所示，本方法的收敛速度要快于其他对比方法，并且最终收敛后的准确率也高于其他方法。可以注意到，在开始阶段本方法的准确率上涨的最快，这是由于一开始模型训练的不充分，稳定精确的信号传输是准确率上涨快慢的关键因素。由于本方法相比一跳传输的方法(FedAvg，FedAvg\_Aircomp)设置了中继，信道条件更好，因此在联邦学习的模型上传阶段信号传输的更准确，在训练过程中收敛速度快，准确率提高得将更快。另外，相比于其他两跳传输的方法(FedAvg\_Dis)，本方法在分簇时考虑了模型的可解释性，利用该解释分析模型推理的内在原理得到该模型的可解释性矩阵，将该矩阵作为指标实现模型的分簇，实现了更优的聚类方案，因此最终的收敛速度与最终准确率将优于对比方案。

图4‑2 收敛速度对比

### MSE

图4-3展示终端数量对各节点的平均MSE的影响。如图所示，本方法的的平均MSE相比其他对比方案更小，并且随着终端数量的增加，各方法的平均MSE也随着减小。不难看出，应用空中计算的一跳传输的方法(FedAvg\_comp)的MSE明显大于其他的方法。这是因为空中计算技术将信号在空中汇聚，如果出现部分终端到边缘服务器的信道条件差的情况，最终汇聚信号将会显著收到这些差信道的影响，因此其平均MSE会大于其他方法。因此通过设置中继来改善信道条件以提高MSE是有效的办法。

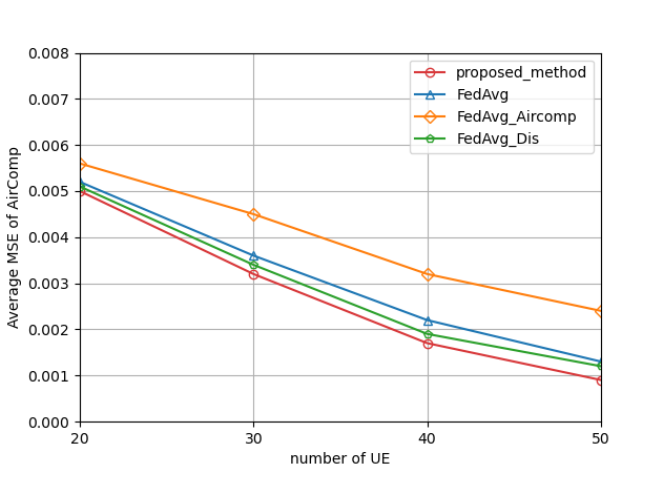


图4‑3 节点数量对MSE的影响

图4-4展示噪声功率对平均MSE的影响。如图所示，本方法的的平均MSE相比其他对比方案更小，并且随着噪声功率的增加，各方法的平均MSE也随着增大。此外，当噪声功率增大时，各算法的平均MSE的斜率也随着增大，并且算法间平均MSE之间的差距也逐渐变大。因此，在噪声功率大的环境下，本算法将展示出更突出的优势。

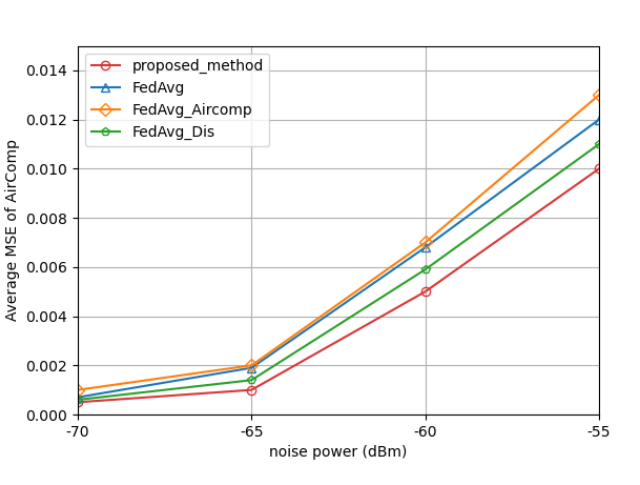


图4‑4 噪声大小对MSE的影响

本章还对比了所提算法与不通过功率控制的算法之间的MSE，从而探究功率控制对于最终信号MSE的影响。图4-5展示了不同终端数量下功率控制算法对于MSE的影响，其中不进行功率控制的算法均使用功率均为最大功率。由图4-5所示，使用功率控制的算法MSE要低于全功率输出的算法，并且两跳传输算法的MSE会小于一跳传输算法的MSE。结合4.3.2的理论证明，不难发现，在大数情况下设置了中继的两跳信号传输的信号失真率要低于一跳信号传输的信号失真率。此外，随着终端节点数量的增多，两跳传输与一跳传输的信号失真率相差越来越大。这是因为终端节点数量越多，终端与边缘服务器之间的信道条件更加不确定，出现部分信道差的情况增加，因此最终影响到最终的MSE。

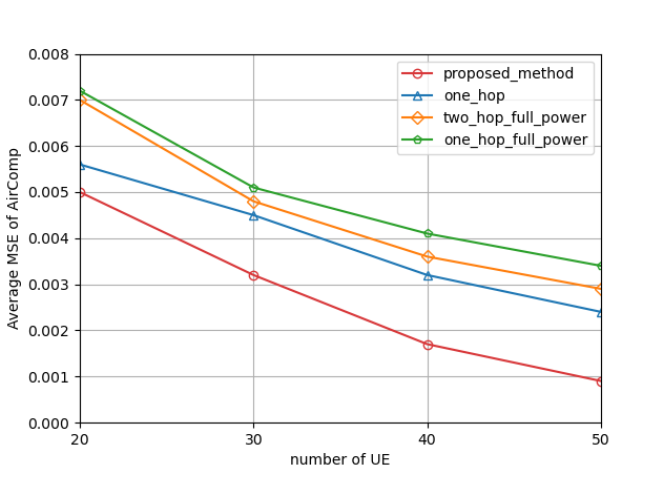


图4‑5 不同功率控制算法的MSE

### 准确率

表4-6展示本方法与其他基准方案的准确率对比。如表所示，在不同终端节点数量的情况下，本方案在准确率均高于其他对比方案，并且随着终端节点数量的增大，各算法的准确率也随着增大。其中由于FedAvg\_Aircomp 为一跳传输，并且没有设置中继来保障良好的信道条件，因此信号的失真率会相对更大，所以最终的准确率相对较低。FedAvg没有应用空中计算技术，部分终端到边缘服务器的信道条件较差不会直接影响到聚合模型，因此信号失真率对于准确率的影响相对会更小。FedAvg\_Dis是完全依赖距离来进行分簇的，只从考虑了信道因素，没有关注模型之间的关联，因此在MSE与本方法相当的情况下，准确率比本方法低。

表4-6 准确率对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 节点数 | 20 | 30 | 40 | 50 |
| FedAvg | 91.375% | 91.575% | 91.625% | 92.325% |
| FedAvg\_Aircomp | 87.125% | 87.875% | 87.985% | 88.875% |
| FedAvg\_Dis | 91.675% | 91.575% | 91.925% | 92.625% |
| Proposed\_method | 92.125% | 92.265% | 92.375% | 93.125% |

## 本章小结

为了应对分布式机器学习存在通信频域资源有限，带宽受限的问题，本文研究了基于空中计算的可解释性联邦学习中功率控制算法。考虑到终端到边缘服务器的信道条件差，会明显影响信号传输的失真率，本章首先提出了设置中继节点的多层次联邦学习架构，并构建边端协同场景下基于空中计算的两跳信号传输数学模型，然后提出了一种基于迭代的算法来求解问题。此外，本章还提出了一种基于可解释性的模型聚类方法，实验显示该方法比基于距离的分簇方法效果更优。仿真结果表明，与其他联邦学习算法相比，本算法在收敛速度、MSE与推理准确率都优于其他对比算法。并且，从理论推理与仿真的角度同时说明了两跳信号传输的MSE将低于一跳信号传输的MSE。