超蜂窝网络柔性覆盖与控制

毕文平¹, 徐修强³, 高原¹, 陈国峰¹, 陈雁³, 周世东^{1,2*}

1. 清华大学电子工程系, 北京 100084

2. 清华大学信息科学与技术国家实验室(筹), 北京 100084

3. 华为技术有限公司, 上海 450008

* 通信作者. E-mail: zhousd@tsinghua.edu.cn

收稿日期: 2017-02-15; 接受日期: 2017-03-23; 网络出版日期: 2017-05-09

国家重点基础研究发展计划(973)(批准号: 2012CB316002)、港澳台科技合作专项项目(批准号: 2014DFT10320)、国家高技术研究发展计划(863)(批准号: 2015AA016202)、清华大学自主科研项目(批准号: 2016ZH02-3)、国家科技重大专项项目(批准号: 2015ZX03001031-002) 和 华为公司创新研究计划资助项目

摘要 面向绿色移动通信而提出的超蜂窝网络架构, 是建立在以业务和控制覆盖分离为核心的柔性覆盖基础上的, 本文重点研究了控制与业务覆盖的分离覆盖架构, 给出了分离覆盖的原则以及分离覆盖的条件和实例, 并在此基础上, 就控制过程的两个核心问题展开研究. 针对部分信道观察下的信道预估问题, 结合基于环境学习的信道预测机制, 利用信道的时间和空间相关性, 分别提出了基于业务基站周期开启测量进行测量与服务决策, 以及基于离线信道学习和在线部分观察的信道空间预估方法. 针对控制覆盖的大容量接入问题, 通过将稀疏码多址(SCMA)非正交多址接入用于非授权接入技术的方案, 并进一步改进发送和接收方法, 可有效支持控制覆盖所需的大容量、低时延、高可靠的短包控制接入.

关键词 柔性覆盖, 覆盖分离, 信道估计, 非授权接入, 稀疏编码多址

1 引言

未来移动业务需求的发展趋势包括业务需求多样化、业务需求空时不均匀且动态变化、大动态的速率、实时性、业务量需求等^[1]. 为此运营商不得不按最高速率要求去部署基站. 在单站峰值速率受带宽限制难以显著增加的情况下, 只能靠增加基站密度来解决, 然而大部分时间空闲的基站为了保持工作状态, 系统总能耗最终随基站密度呈线性增加, 仅从系统能耗角度看就是难以持续发展的^[2~4].

为了保证移动网络的可持续发展, 我们观察到: 大规模增加基站密度的结果使得处于业务服务状态的基站数量(由整体用户数决定)并不会随之同比增加, 而处于非服务状态的基站比例却在大幅度增加. 这些处于非服务状态的基站, 如果保持活跃状态也要消耗大量电能. 因此及时关闭大部分非服

引用格式: 毕文平, 徐修强, 高原, 等. 超蜂窝网络柔性覆盖与控制. 中国科学: 信息科学, 2017, 47: 544-559, doi: 10.1360/N112017-00042
Bi W P, Xu X Q, Gao Y, et al. Agile coverage and control for hyper-cellular networks (in Chinese). Sci Sin Inform, 2017, 47: 544-559, doi: 10.1360/N112017-00042

务状态的基站,成为解决业务需求导致的网络可持续发展(从绿色节能角度)的必由之路^[5,6].因此牛志升等提出了超蜂窝架构理论^[7].

超蜂窝架构中,基站是随着业务的需求而按需开启和关闭的.非服务状态基站的关闭与开启,必须是动态的、按需的柔性的覆盖^[8].要做到这一点,首先就要对空中传输的不同数据的无线覆盖进行一定的分离,从总体上讲,就是业务覆盖和控制覆盖的分离.将基本的控制数据从业务覆盖中分离出去,可为动态关闭业务覆盖基站提供基本的条件.在此基础上,还要解决两个层次的问题:一个是控制过程,另一个是控制算法.从控制过程看,就是要有一个控制的层面和实体,让控制者(网络侧的控制器)可以获得控制决策所需的上下文信息,同时要把控制者的决策及时准确地传达到执行者(终端、基站、RRU(远程射频单元 Remote Radio Unit)、交换机等).控制算法的作用,则是让控制者能根据所得的上下文信息,得到网络中用于控制各个实体单元的参数(包括终端、天线、资源分配、开关等).在给出超蜂窝网络的覆盖架构和准则,并提供了控制方式,获得信道预测信息后,需要对系统的能效以及谱效进行优化,方法包括小区开关^[9]、大规模 MIMO^[10] 和全双工^[11] 等方法.

根据以上内容的讨论,关于超蜂窝网络架构中的柔性覆盖需要开展以下 3 方面的工作:(1) 覆盖分离是动态开关的使能机制,为此,需要开展第一个层面——控制覆盖与业务覆盖分离机制及优化的研究.(2) 控制过程是分离覆盖后动态开关的使能环节,为此需要开展第二个层面——超蜂窝网络中控制覆盖的能效优化方法的研究,重点在解决业务基站关闭与其信道状态信息获取之间的矛盾,以及控制基站覆盖下海量潜在用户的接入问题.(3) 在网络侧获取信道信息之后,剩下的就是进一步优化系统的效率问题了,为此需要开展第三个层面的研究——超蜂窝网络动态小区形成方法研究.

本文将对第一、二方面工作内容重点介绍.主要包括以下 3 部分内容.

A. 控制覆盖与业务覆盖分离机制及优化研究. 在控制覆盖与业务覆盖分离机制方面澄清了分离覆盖的基本属性:即为完成一件任务而允许不同的数据经过不同的基站设备提供覆盖.在由大小两层覆盖组成狭义的分离覆盖中,明确了大覆盖除了支持控制覆盖以外,也允许支持低速业务覆盖,而小覆盖除了支持高速业务覆盖以外,也允许支持控制覆盖.提出了广义分离覆盖的概念,将控制信令也当成空口多种数据覆盖的一部分,支持多层覆盖,根据各种不同数据的传输要求,确定为其覆盖的层.新的系统架构可实现向下兼容和平滑演进,广义分离可后向兼容狭义分离,分离覆盖可兼容现有蜂窝架构.提出覆盖层与设备解耦的概念,即允许同一设备可以在不同时刻提供不同的覆盖层服务.提出上下行覆盖解耦的概念,即为同一用户提供下行数据发送和上行数据接收服务的基站设备可以在空间上分离.上、下行覆盖可以独立设计,提供非对称服务,具有很高的设计灵活性.

B. 超蜂窝网络中控制覆盖的能效优化方法. 在控制覆盖的探测方面,论文中将介绍一套基于部分信道观察的信道预估方法.在超蜂窝网络中,从控制覆盖中分离出来的业务基站,被要求必须能关断以节能.然而业务基站关闭后,随之也关闭了它的导频发射和(或)接收测量的能力.在失去这种对信道的直接观察时,如何解决业务基站重新开启的决策依据来源,是超蜂窝网络不可避免的新问题.论文介绍的信道预估方法特点主要体现在:基于环境学习的信道预测机制,通过分析与测量,证明了信道的空间和时间上的相关性是部分观察下获取信道信息的重要而有效的途径.利用信道的空间相关性,提出了基于业务基站周期开启进行测量与服务决策的机制.利用信道的空间相关性的基于离线信道学习和在线部分观察的信道预估方法.分别针对直接信道预估和借助于用户定位的间接信道预估提出了相应的方法.

C. 采用 SCMA 实现非授权接入技术进行控制覆盖. 在控制基站覆盖下海量潜在用户的接入方面,针对现有 LTE 网络中广泛使用的数据传输方法是基于请求、调度模式,不仅消耗大量的额外无线资源,也引入了比较大的控制延时.本文提出将免授权接入机制引入控制覆盖的业务请求,通过随

时接入, 可以大大减少接入延时和信令资源开销比例. 进一步引入稀疏码多址信号设计, 引入随机接入, 使得在相同无线资源下, 多个用户同时接入时, 基站仍可以分辨并成功解码. 相对于其他多址方法不仅可以支持多用户接入, 还可获得一定的编码增益. 更多用户同时接入能力与更高的超载率相结合, 可使得在免授权接入中, 获得数倍的频谱效率提升, 可显著提高一次传输成功率, 实现高可靠、低延时的目标. 在这方面本文将重点介绍在控制覆盖中应用免授权接入、稀疏编码多址的方法及相应的适应性改进方法.

本文结构安排如下: 第 2 节介绍对控制覆盖与业务覆盖分离机制及优化. 第 3 节介绍部分信道观察下的信道预估方法的主要思路. 第 4 节介绍蜂窝网络的控制覆盖的非授权随机接入相关技术. 第 5 节为本文结论.

2 控制覆盖与业务覆盖分离机制及优化

2.1 广义的分离覆盖

在现有蜂窝体系结构及其发展过程中, 都是以基站为中心的, 由于其多重身份 (测量、控制、传输等), 占据着异常重要的角色. 然而随着当前及未来业务需求多样性的呈现以及空间、时间业务不均匀性的突显, 在未来高效绿色通信的约束下, 无法通过关闭大量低业务或无业务基站以实现可持续发展. 本文提出了以控制覆盖与业务覆盖分离的基本思想为核心的超蜂窝架构, 将无处不在的、永远在线的控制覆盖与按需开启的动态业务覆盖相结合, 可实现控制覆盖和业务覆盖之间相对独立的优化和演进, 代表了新一代移动通信网络架构的新趋势. 相对于现有蜂窝结构面向峰值需求规划的覆盖方式 (这里称为硬性覆盖), 给出柔性覆盖的定义: 根据系统的业务需求和网络的实时状态, 为用户提供动态的按需覆盖, 以达到提升能效和谱效的目的, 支撑未来业务的可持续发展^[8].

从能效的角度来看, 不管是信令还是业务, 它都是属于需要承载在空中信号上的数据, 我们要做的是关注不同的数据该如何更好地进行映射, 而不同的映射显然不能简单的根据它是用于业务还是控制进行区分, 应该考虑数据的本质特征以及与其他业务的相关性, 综合得出最优的分层方案. 这些特征应包括数据量、时间到达规律、时延要求、数据包之间的相关性等.

因此分层超蜂窝覆盖作为一种新型架构, 狭义上讲, 是控制和业务覆盖的分离, 而广义上讲是多层数据覆盖, 即未来业务服务过程中所需的覆盖数据 (包括用于控制的数据和用户业务的数据), 均可根据网络服务状态, 按需地进行分离覆盖, 通过多层覆盖的协作, 完成高效、高质量的覆盖服务.

柔性覆盖的核心思想是控制和业务的分离, 但在具体架构上又不拘泥于控制与业务的分离. 为此, 先对一些概念进行定义:

- **控制覆盖.** 以控制数据为主的数据覆盖, 同时可以支持一定条件的业务数据.
- **业务覆盖.** 以高速业务数据为主的数据覆盖, 同时可以支持一定条件的控制数据.

下面以双层覆盖为例说明^[12]. 该双层覆盖无线网络由控制覆盖和业务覆盖两层组成, 其中, 控制覆盖是指拥有较大覆盖范围, 主要传输控制数据, 在资源允许的情况下兼顾数据业务的覆盖. 业务数据覆盖是指拥有相对较小的覆盖范围, 主要提供高速数据业务数据的覆盖, 同时兼顾数据传输过程中的控制.

为了清晰地描述该架构的特点, 图 1 给出了该架构的示意.

此外, 在双层覆盖网络的帮助下, 尽管某些业务覆盖节点处于关闭或休眠状态, 该节点下用户的接入请求可以由控制覆盖进行感知, 再由网络侧决策并通知相应的数据覆盖开启, 以完成对应的服务,

图1 双层覆盖网络架构示意图
Figure 1 Two-layer coverage structure

从而解决休眠基站“只能关不能开”的难题。

2.2 分离覆盖中的数据映射原则

数据传输需求的主要特征包括: 延时要求、速率要求、持续时间 (针对数据流或周期性数据)、到达率需求、允许的接入等待时间等。系统应根据这些需求特征, 确定每一次通信过程中涉及到的各阶段、各类型数据的缺省 (初始) 映射, 即在不考虑网络状态情况下的初步映射决策。进而, 当网络状态发生变化时, 具体的实际映射, 也应随之作一定的调整。

从总体上看, 分层和映射的大体原则应该是:

- (1) 高速业务数据映射到小覆盖层, 作为缺省映射层。
- (2) 低速业务数据, 特别是高实时性的低速业务数据映射到大覆盖层, 作为缺省映射层。
- (3) 数据到覆盖的映射可以根据不同的方案动态调整, 即层间数据迁移。

2.3 数据覆盖的迁移原则

在双层覆盖中, 控制覆盖采用的是宏基站大面积覆盖, 而业务覆盖由相对密集的小区或基站组成。我们也需要注意到, 业务基站既然可以支持高数据率的数据传输, 自然本身也有能力支持低速率数据的传输。因此当业务基站开启时, 它在服务高速率数据传输的同时, 还可以分担一部分它的覆盖范围内的低速率数据。

因此在一定条件下, 应允许一部分低速业务迁移到业务覆盖里去。这种迁移条件包括:

- (1) 与激活业务覆盖具有很强的时间、空间相关性的数据 (包括控制数据和业务数据)。
- (2) 在一段时间内, 某个非激活业务基站覆盖区 (覆盖空洞) 中集中出现大量的低速业务 (时间、空间聚焦)。
- (3) 面向移动性支持的数据, 不参与向业务覆盖迁移。
- (4) 高速移动用户的数据, 在小区边界之前, 应迁移到控制覆盖以保持连接。

图 2 数据覆盖中的分离覆盖及 offloading 示意图
Figure 2 Decoupled coverage and offloading in data coverage

图 2 给出了数据覆盖中的分离覆盖及数据卸载示意图.

2.4 分离实例

本文基于数据信令分离的超蜂窝网络架构, 针对 LTE 的演进, 具体研究了实现异构网络数据和信令的解耦合的可行途径, 具体分离细节详见文献 [13, 14].

2.5 面向高能效的新型蜂窝网络架构的特点

这里总结一下新型的超蜂窝网络架构的特点, 它由至少包括控制覆盖和业务覆盖的多层覆盖组成. 该架构的核心特征包括:

- (1) 超蜂窝组网和覆盖, 首先要有多重的覆盖 (可以是同体制的多重覆盖, 也可以是不同体制的, 可以是同频段的, 也可以是不同频段的).
- (2) 允许覆盖空洞: 某些层的覆盖 (业务覆盖) 可以动态关闭, 以节能.
- (3) 一个通信任务允许通过多层覆盖来完成 (不同的信令, 不同的业务数据).
- (4) 多层覆盖之间的设备具有协同能力, 这是完成通信任务的必须要件, 也是优化的前提.
- (5) 可软件定义的设备平台, 提供支持多覆盖的设备共享, 传输覆盖的灵活升级, 更重要的多覆盖之间的协作.

超蜂窝架构除了上述这些基本特征之外, 在具体的架构实现上应该是开放的, 可有多种选择, 不断演进. 包括:

- (1) 覆盖设备间及与控制设备间的网络拓扑.

(2) 可软件定义的公共设备平台的软件和硬件架构.

(3) 具体的控制优化方法.

为了实现这种协同的多重覆盖架构,不同的演进阶段,可以通过多种设备形态来完成,至少包含两类设备形态:每层覆盖有自己的系统和设备;各层覆盖共享设备.

在第1类设备状态下(每层覆盖有自己的系统和设备),很多设备是跟天线捆绑的,也包括关闭这些相关设备,关闭天线与射频往往也等价于关闭对应的基站设备,从而降低基本处理能耗.

而在第2类设备状态下(各层覆盖共享设备),由于公共平台本身不能整体关闭,而涉及的覆盖天线与射频是一对一的,同时射频部分也是能耗的最主要部分(随着天线分布密度增加,射频数也在同步增加),因此射频部分的关闭是节能的最主要环节.由于采用公共平台,分布式无线通信系统(DWCS),所有天线(或大部分密集部署的天线)共享同一处理平台和入网接口,为了节能,就要求公共处理平台具有以下的特点:

(1) 硬件处理设备与天线解耦,这部分功能可以通过前传交换网络来实现.

(2) 处理运算只与用户业务量或活跃天线数成正比,与总天线数量无关.

(3) 处理能耗只与运算量成正比.

这样的天线与硬件处理解耦的方式,除了可以提供设备形态之外,也为未来引入开放的软件定义无线网络和无线网络功能虚拟化,提供了一个可以切片化定制服务的平台,以适应5G网络垂直应用的需求.

3 用户位置及与业务基站间信道信息获取

在柔性覆盖中重启关闭(唤醒休眠)基站时,本应依据该基站信道特性作决策,然而由于该站的关闭,无法用传统方法获得它与目标用户之间的信道信息.

为了获得信道信息,就必须有一定的测量与信道估计.这里信道估计的首要目的,是用于确定开启哪些基站,特别是尚未开启的基站,为当前的用户或用户集合服务.这里的信道估计又可以分为基于如GPS、惯性导航等的额外位置信息^[15]和不基于额外位置信息两大类方法^[16~19].

对于基于非额外位置信息的方法,可以分成两种情形:利用活跃基站进行信道测量^[16,17]和开启非活跃业务基站进行信道测量^[20].开启非活跃业务基站进行信道测量,又可以分成基站侧发导频,终端侧发导频,两侧互发导频3种.对于以选择唤醒基站为目的信道测量而言,精度要求并不高,因此两侧互发导频的意义不大.剩下两种,从单条链路的测量上看,是基本等价的.只是从功耗上看有一定的差异,终端侧发导频的方案中空闲业务基站可以只保留接收模块的工作,而可以关闭发射模块.基站侧发导频的方式则可以让空闲业务基站关闭接收模块而保留发射模块.确定了由哪一侧发导频以后,为了进一步控制能耗,应选择部分时间,开启部分业务基站进行导频发射或接收.

为此,我们从基于环境学习的信道预测机制、利用信道的时间和时间相关性3个方面开展相关研究.

考虑到在地面蜂窝移动场景中,电磁波传播环境中的大部分地理环境是静止的.信道之所以会变化,主要来源于少量的环境参数(如反射/散射系数等),以及用户本身在地理空间中的移动.为此,基于环境学习的信道预测机制^[21],通过平时积累的测量数据,结合动态的部分测量(部分时间、部分空间上的信道的测量,即部分时间里开启的部分基站所观测到的动态数据)对全时空的信道传输情况进行预估,相当于信道参数在时间、空间上的插值和预测.

图 3 一维 T_2 对小基站总能量消耗 E 的影响 [20]
Figure 3 T_2 vs. total energy consumption E [20]

图 4 二维 T_2 对小基站总能量消耗 E 的影响 [20]
Figure 4 T_2 vs. total energy consumption E [20]

在这个基本思路下, 这里分别介绍基于信道的时延相关性和空间性的两类以选择激活休眠基站为目标的信道预测方法。

3.1 基于时间相关性的方法

此时暂不考虑不同基站的信道之间的空间相关性, 但考虑到用户移动是连续的, 其位置不会在短时间里发生大范围的跳变。因此, 可以让用户对业务基站的信道观测以周期性测量的方式进行。这样业务基站只需要在一个业务周期 T_2 里开启一小段时间 T_1 进行测量 ($T_1 \ll T_2$), 获得用户在此时刻与各业务基站的信道关系并排序, 然后在业务周期的静默期间, 只要有这个用户的业务, 就将之前测得的最好的那个基站开启, 为该用户服务。显然两次开启测量基站的间隔是一个重要参数。测量间隔越大, 测量本身带来的能耗就可以降低, 然而测量间隔越大, 也会带来测量精度的下降, 使激活不合适的业务基站的概率增加, 又会带来业务传输能耗的增加。

开启业务周期的设计是影响该机制实施效果的关键因素。为了使网络总能耗最小, 我们通过理论分析和仿真两种方式进行了最优开启周期的设计 [20]。

在一定的系统参数设置下, 一维和二维空间的仿真结果如图 3 和 4 所示 [20]。

图 3 给出了小基站总的能量消耗 E 随着业务期的持续长度 T_2 的变化关系。图中纵向虚线标识了我们通过理论推导选取的 T_2 , 可以看出其与能量消耗 E 的曲线相交于最低点。 D 为用户运动模型中的参数, 其越大表示用户运动越活跃。对比 D 不同的两条曲线, 可以看出, 用户运动的活跃性越大 (D 越大), 总能耗越大, 而最优的探测周期 T_2 越短。

针对二维场景, 我们结合更一般的路损模型, 引入了阴影衰落以及阴影衰落的空间相关性。结果表明随着 T_2 的增长, 信道信息获取时刻和网络利用此信息做决策时刻的平均时间间隔变大, 导致决策的误差变大 (例如: 选择并开启了一个距离较远的小基站为用户提供服务), 继而用于传输的能耗变大。用户运动的活跃性越大 (D 越大), 业务传输能量消耗 E_2 越大。

3.2 基于空间相关性的方法

本节考虑空间相关性。也就是说, 考虑在需要了解某个时间用户到某个区域内所有基站的相对信道关系时, 只利用该时刻的部分基站进行信道的测量, 然后根据该时间的部分基站信道观察值去估计

图5 间接法

Figure 5 Indirect method

图6 直接法

Figure 6 Direct method

该用户到所有基站的信道参数。

为此, 我们提出了一系列基于环境学习的信道预测和决策的方法^[18,19], 包括间接法和直接法, 如图5和6所示, 它们都包含两个阶段: 离线学习阶段和在线预测阶段. 间接法的基本思路是以位置为中间参量, 利用对用户到部分基站之间的信道观察, 结合历史数据学习, 对用户所处的位置进行估计, 然后再利用用户的位置结合历史数据对信道特性进行估计. 直接估计法的思路则更为简单直接, 即利用历史数据和部分基站的测量, 直接得到所有基站到该用户之间信道的信息.

3.2.1 间接方法中的基于在线学习的用户定位

A. 在线学习中的接入点选择和数据库选择. 传统室内定位方案假设离线和在线阶段环境是稳定的, 在线定位时不考虑在线环境状态, 然而实际环境的变化, 可能会降低定位的精度和稳定性. 为了解决这个问题, 引入接入点 (AP) 选择策略, 根据环境变化动态做出 AP 选择, 并及时地更新数据库提升定位性能^[18].

B. 基于环境学习的指纹数据库更新策略. 数据库的更新或者选择是针对环境的慢速持续变化. 慢速持续变化是经过在线阶段大量统计之后对环境学习得到的结论^[19].

C. 基于历史信息的在线学习定位算法. 既然更新数据库需要大量用户长时间学习, 这些历史信息也可以帮助定位, 使得定位性能优化. 利用基于环境学习的定位机制, 结合历史定位信息, 可以引入用户的移动轨迹修正引导室内位置确定, 给定当前时刻之前一段时间用户所在位置以及运动轨迹, 可以给出用户当前出现在不同定位区域的出现概率, 将其引入传统的定位算法, 利用不同区域的出现概率修正定位结果, 做为环境因素的调整.

3.2.2 基于空间相关性的直接信道预估方法

基于指纹数据库的动态基站开关策略分为单用户实时决策问题和多用户联合决策问题. 单用户实时决策每次选择信号最强的基站服务, 可以分为选择单基站和选择基站集合. 多用户联合决策则要为所有用户选择合适的服务基站集合. 对于单天线用户, 每次只要选择打开一个基站为用户服务, 对于多天线用户可以选择基站集合联合为用户服务.

图 7 基于指纹的基站选择流程^[19]

Figure 7 Fingerprint based BS selection steps^[19]

基于指纹数据库的动态基站开关节能方法包括离线阶段和在线阶段, 在线阶段的核心是动态基站开关决策策略. 离线和在线阶段流程如图 7 所示.

3.3 基于信道测量的基站选择策略分析验证

这里基于如图 8 所示的室内办公室场景信道测量, 对基站选择策略进行方案验证和性能评估^[19].

利用加权 K -近邻 (WKNN) 指纹方法进行动态基站开关策略实验, 分析平均接收功率与基站选择正确率.

实验环境中共有 6 个基站, 在每次决策时根据打开的探测基站个数 p 的不同, 基站选择正确率如表 1 所示. 其中, $p = 0$ 时随机选择基站, $p = 6$ 时可以直接选择 RSS 最大的基站, 选择正确指的是选中实际信号 RSS 最强的基站.

4 控制覆盖中的免授权接入技术

4.1 永远在线的控制覆盖关键技术 —— 免授权随机接入

为了保证柔性业务覆盖可能带来的暂时覆盖空洞中的用户 (暂时没有业务) 在发生业务需求时能得到超蜂窝网络的及时服务, 需要控制覆盖提供永远在线的接入申请服务. 事实上, 只要网络侧可以及时得知哪些用户需要业务覆盖服务, 即可通过启动相应的探测过程了解到该用户位置和信道信息, 进一步确定所需开启的基站或扇区, 同时分配相应的时频资源、天线波束, 从而提供按需的柔性覆盖. 因此网络侧对用户业务发生的及时感知是柔性覆盖的关键. 而这个感知的基础就是用户业务服务请求信息的及时上传, 这类请求信息往往只需要包括用户身份、终端状态、请求的业务种类及可能的少量参

图 8 基于信道指纹的基站选择实测场景平面图
Figure 8 The actual BS deployment of fingerprint based BS selection

表 1 基站选择正确率
Table 1 The accuracy of BS selection

p	0	1	2	3	4	5	6
Accuracy rate	16.67%	59.02%	80.33%	85.25%	96.72%	96.72%	100%

数, 因此属于非常短的消息, 但其传输的实时性和可靠性要求非常高. 这是控制覆盖必须要首要支持的数据传输服务.

现有蜂窝网络协议, 在有数据业务时, 需要先执行一个数十毫秒到上百毫秒的连接建立过程. 而对于超蜂窝网络中控制覆盖而言, 对某个具体的用户而言, 其实很可能数据量很少, 却占用了大量的控制信道资源, 且引入了相当大的时延 [22].

我们知道, 4G 网络中广泛使用的数据传输方法是基于请求、调度模式的, 这种传输机制是一种基于授权的接入机制. 可是这种传输技术并没有对数据包的大小或者类型加以区分, 因此这种传输机制无法用于信令网络传输极大量的低时延、高可靠的小包数据 [23].

为此, 我们提出针对此类极大量、低时延、高可靠要求的小数据包的传输, 采用免授权机制的方法. 从资源划分上, 系统将预留一部分时频资源供免授权短包接入使用. 经过初始化的移动设备, 只要有此类小包业务需要发送, 则随时接入, 即随时在免授权接入资源中发送待传的数据短包, 而不事先发送任务请求 [24].

由于免授权接入中, 不可避免地会产生资源的碰撞, 因此, 支持多用户解码的信号设计非常重要. 需要在使用尽可能少的资源开销情况下, 提供对尽可能多用户的识别, 以及多用户资源碰撞时尽可能高的数据解码成功概率.

为了定量评价免授权机制的性能, 我们定义一些评价指标和系统参数. 评价指标主要是资源利用率, 可以用平均每资源可支持的平均用户数来衡量, 它与系统指标参数, 如可支持的服务用户数有一

图 9 净吞吐量与总到达率之间的关系 $f_K(\lambda)$
Figure 9 Throughput vs. arrival rate

定的换算关系, 以资源率、每用户数据包到达率为参量.

当小区短包用户数为 N , 每用户业务初始发起 (不算重传的) 平均时间间隔为 T , 单位时间里可用于免授权接入的时频资源数为 M , 每个消息包传输所占时频资源数为 m , 每个用户所允许的接入时隙间隔为 t_1 , 平均每次重传延时时为 t_2 , 假设每用户业务发起为独立的 Poisson 过程.

平均每个消息传输资源包的时间间隔为 m/M , 因此每个用户在每个平均传输间隔内, 有业务初始发起的概率为 $p_0 = \frac{m}{MT}$. 而每次有机会传输时, 原始信息包的平均数量为 $\lambda_0 = \frac{mN}{MT}$, 因此希望网络在免授权竞争接入和重传的过程中, 达到净吞吐率 R (定义为平均每个传输资源包可实现的正确传输的消息包数量), $R = \lambda_0 = \frac{mN}{MT}$.

假设采取某种重传方法, 使得整体看起来仍近似于各用户呈独立的 Poisson 接入过程. 由于存在重传, 每个用户在平均每个消息传输资源包中可能要传输的概率为 $p > p_0$. 在每个消息传输时频资源包需要发送数据的用户数期望 (这里称为总到达率) 为 $\lambda = pN$. 出现 k 个用户同时接入的概率记为 P_k , 当 N 非常大, k 很小, p 很小时, 可以证明: $P_k = \frac{\lambda^k}{k!} e^{-\lambda}$. 当用户采用的多址接入方法在 k 个用户同时接入时的误包率记为 P_{ek} (这里假设了各用户信道条件相当的条件), 则每次成功传输的吞吐量 (相当于净吞吐量 λ_0 或 R) 与总到达率 λ 之间的关系为

$$R = \sum_{k=1}^N P_k k (1 - P_{ek}) = \sum_{k=1}^N (1 - P_{ek}) \frac{\lambda^k e^{-\lambda}}{(k-1)!} = g(\lambda), \quad (1)$$

注意这里的函数 g 与多址方式下的多用户译码能力 P_{ek} 有关.

下面考虑一种近似的情况. 在很多情况下, 误包率与碰撞用户数 k 之间的关系呈现明显的台阶形状, 即最多可同时分辨 K 个用户. 当 k 不大于 K 时, P_{ek} 近似为 0. $k > K$ 时, P_{ek} 近似为 1. 此时成功传输的吞吐量 (净吞吐量) 与总到达率 λ 之间的关系为

$$R = \sum_{k=1}^K k P_k = \sum_{k=1}^K \frac{\lambda^k e^{-\lambda}}{(k-1)!} = f_K(\lambda). \quad (2)$$

不同的可同时分辨用户数 K 对应的净吞吐量与总到达率之间的函数曲线如图 9 所示.

图 10 最优净吞吐量与可分辨用户数的关系

Figure 10 Throughput vs. number of recognizable users

可以求出处于稳态时的含重传的总到达率 $\lambda_1 = f_K^{-1}(\lambda_0)^{-1}$. 据此, 可以得到此时平均每个有数据要传的用户的一次传输成功率为

$$P_{\text{suc},u} = \frac{\lambda_0}{\lambda_1} = \frac{\lambda_0}{f_K^{-1}(\lambda_0)}. \quad (3)$$

每个数据包的平均传输次数为

$$\bar{L} = \frac{1}{P_{\text{suc},u}} = \frac{f_K^{-1}(\lambda_0)}{\lambda_0}. \quad (4)$$

因此, 总延时中, 除了一次接入占用的时频资源包的时长之外, 额外延时为

$$\tau_{\text{ext}} = \frac{t_1}{2} + (\bar{L} - 1)t_2 = \frac{t_1}{2} + t_2 \left(\frac{f_K^{-1}(\lambda_0)}{\lambda_0} - 1 \right). \quad (5)$$

从图 10 中可以看出, 随着可同时分辨的用户数从 1 增加到 6, 对应于最优净吞吐率的平均传输次数也从 2.7 下降到 1.4, 即占用的控制信道资源数下降约到原来的 1/2, 相应的重传次数从 1.7 次下降到 0.4 次, 对应于总延时中的主要部分——重传延时下降到原来 1/4.

这样, 具体的评价针对给定的每用户业务初始发起 (不算重传的) 平均时间间隔 T , 单位时间里可用于免授权接入的时频资源数 M , 问题包括:

(1) 给定当小区短包用户数 N , 通过设计不同的多用户编码方案 (决定每次传输要用的资源数量 m), 评价其重传次数或时延.

(2) 给定重传次数要求或时延要求, 通过设计不同的多用户编码方案 (决定每次传输要用的资源数量 m), 评价可支持的总用户数 N .

对 (1), 评价方法就是, 根据给定的 N 和设计编码得到的 m , 求得所需达到的净吞吐率 $\lambda_0 = \frac{mN}{MT}$, 然后根据编码设计得到可支持的最大同时用户数 K , 利用图 9 的 $f_K(\lambda)$ 关系求得总吞吐量 λ , 从而得到 $\bar{L} = \frac{1}{P_{\text{suc},u}} = \frac{f_K^{-1}(\lambda_0)}{\lambda_0}$.

1) 注意, $f_K(\lambda)$ 是一个单峰的拟凹函数, 因此其反函数最多有 2 个解, 我们关心的是最小的正解.

对 (2), 对于给定的时延要求, 即平均重传次数, 可在图 9 中做出斜率为平均传输次数倒数 $1/\bar{L}$ 的过原点的直线, 与根据编码设计得到的可支持最大同时用户数 K 对应的曲线交点的纵坐标即为该种编码方式下可支持的净吞吐量, 进而利用编码设计得到 m , 计算可支持的总用户数 $N = \frac{\lambda_0}{m} MT$. 其中 MT 的物理意义是单个用户平均业务时间间隔内系统可用的总的资源数. 因此可以看出, 系数 $\frac{\lambda_0}{m}$ 是单位无线信道资源可支持的总用户数, 是这种随机接入方式下的频谱效率的一种表征.

可见, 为了达到尽可能高的用户容量, 需要使得满足时延要求下的 $\frac{\lambda_0}{m}$ 越大越好. 我们又可以把它写成如下的形式: $\frac{\lambda_0}{m} = \frac{K}{m} \frac{\lambda_0}{K}$. 其中前一个因子 $\frac{K}{m}$ 体现的是多用户编码带来的同时支持用户数增益与可能需要扩张的资源代价之间的比例, 由编码方法决定. 后者 $\frac{\lambda_0}{K}$ 则完全由图 9 来决定.

综上, 免授权随机接入的多用户编码设计要点就是:

- (1) 增加可同时分辨用户数 K , 从而获得尽可能高的统计复用增益 $\frac{\lambda_0}{K}$.
- (2) 尽可能用较小的资源扩展, 获得尽可能多的同时可分辨用户数, 或在增加同时用户数的同时提高超载增益 $\frac{K}{m}$.

4.2 基于稀疏码多址接入的上行免授权传输

如前所述, 在超蜂窝网络架构下, 高效、实时的控制覆盖是必要的支撑条件, 而免授权接入是完成此项任务的重要手段, 可以避免在短包数据接入中传统的基于连接的接入手段带来的相对巨大的控制开销. 而在免授权接入中, 如前一节的分析所述, 为了取得更高的用户容量, 接入信号设计是非常关键的环节, 为了支持多个用户在相同的时频资源上传输, 必须采用一定的非正交多址技术, 传统的直接序列扩频 CDMA 技术就属于这一类. 不过它获得的可同时分辨用户数 K 的增加, 是以带宽扩展为代价的, 且扩频比要大于可支持的同时用户数, 即超载增益 K/m 一般要远小于 1. 这使得系统容量将大大受限.

近年人们通过一定程度上引入编码, 来实现多用户编码传输, 通过提高超载增益均有希望成为未来的免授权接入的编码调制方案. 包括交织多址^[25], PDMA, MUSA, SCMA 等. 这里以 SCMA 为例介绍其在控制覆盖的免授权接入中的可能发挥的作用.

稀疏码多址 (SCMA)^[22, 23, 26~29] 是一种非正交多址技术, 也是一种有效的多用户编码技术, 它通过在多个可用调制资源里选择部分资源结合多维调制编码, 形成不同的 SCMA 码本, 可实现超过资源占用数的可分辨用户数, 即获得大于 1 的超载能力. 因此非常适合用于超蜂窝控制覆盖的上行免授权传输. 基于稀疏码多址接入的上行免授权传输技术中用户发送上行数据包流程如下: 随机或根据特定的规则选择 SCMA 码本, 使用所选择的 SCMA 码本 (SCMA 码本与导频序列有映射关系) 处理待发送数据包. 数据包处理完成后, 用户直接在约定的上行资源 (如时频资源) 上发送数据包, 而不需要事先向基站发送业务请求并且等待基站的调度指令. 如果有多个用户同时发送数据包时, 基站侧接收到多个用户的叠加信号, 基站基于先进的多用户联合检测接收机, 通过检测导频序列, 可以确定哪些 SCMA 码本被使用, 再借助先进的多用户联合检测接收机, 如消息传递算法, 可以正确解调出多个用户的数据包. 在此过程中, 不同的用户可能会选择相同的 SCMA 码本, 产生码本的碰撞, 但只要导频序列不碰撞, 接收机都能够通过联合检测解调出用户数据.

从具体定量角度来看, 在常见的一种参数配置下, 当 SCMA 码块占用 4 个无线时频资源时, 可以提供 6 个不同的码本, 而且即使多个用户使用相同的码本, 只要它们的信道特性有所不同, 且码本碰撞的用户数不多, 仍可通过一定的接收算法实现多用户数据分离, 大量仿真结果已表明, 在随机接入时, 此参数配置可支持多达 6 个同时用户的数据分离^[23], 因此超载增益 K/m 可以达到 1.5 以上. 与 CDMA 相比, 可以大大提高系统的短包接入容量.

可见免授权传输在小数据包场景下,可以极大地降低系统的信令和时延开销,而稀疏码多址接入又能很好地支持过载系统,支持海量连接.

5 结论

本文针对可实现高能效绿色移动通信的超蜂窝架构进行了丰富和补充. 本文澄清了分离覆盖的基本属性: 即为完成一件任务而允许不同的数据经过不同的基站设备提供覆盖, 并提出了广义分离覆盖的概念, 可兼容狭义分离覆盖及现有蜂窝网络架构. 进一步提出了覆盖层与设备解耦以及上下行覆盖解耦的架构形式, 形成了超蜂窝网络的基本架构特色. 并针对 LTE 向超蜂窝演进中不同覆盖承担信令分解给出了具体的设计方法, 可解决信令复杂度和开销大、切换开销大等瓶颈问题. 在分离覆盖中, 控制覆盖最重要的任务就是: 感知用户业务需求和感知用户与休眠基站之间的信道. 为此, 本文针对业务基站关闭后动态控制所需的信道信息获取问题, 提出通过用户位置作为纽带, 建立起了基站间信道参数之间的间接相关性. 据此, 我们分别提出了利用离线或历史数据库的空间插值方法, 包括间接方法和直接方法. 解决了信道信息获取问题, 并通过实际场景测试验证了基站选择的准确性. 进一步, 为了让系统能及时感知到用户的业务需求, 提出了在控制覆盖中采用免授权接入的方式传输用户的业务请求相关的短消息包, 通过分析验证了在这种大容量、低延时、短包接入需求下, 多用户可分辨能力对系统容量、延时的影响. 指出基于稀疏码多址的免授权接入是一种非常有限的控制覆盖接入手段. 在上述控制与业务覆盖分离的架构下, 在业务感知和信道感知的支撑下, 超蜂窝系统就可以动态地调用各种可能的业务覆盖手段, 进行包括能效、谱效等方面在内的全面优化. 为休眠基站的动态激活、分布式天线的动态协作传输、大规模 MIMO 的动态优化、全双式组网的动态设计等提供基本的系统架构和决策依据.

参考文献

- 1 Meywerk F. The mobile broadband vision — how to make LTE a success. In: Proceedings of LTE World Summit, London, 2008
- 2 Richter F, Fehske A J, Fettweis G P. Energy efficiency aspects of base station deployment strategies for cellular networks. In: Proceedings of the 70th IEEE Vehicular Technology Conference Fall (VTC 2009-Fall), Anchorage, 2009. 1–5
- 3 Griffiths M. ICT and CO₂ emissions. Parliament Office Sci Tech Post-note, 2008, 319: 1–4
- 4 Auer G, Blume O, Giannini V, et al. D2. 3: energy efficiency analysis of the reference systems, areas of improvements and target breakdown. EARTH, 2010, 1–69
- 5 Chen Y, Zhang S, Xu S, et al. Fundamental trade-offs on green wireless networks. IEEE Commun Mag, 2011, 49: 30–37
- 6 Gruber M, Blume O, Ferling D, et al. EARTH — energy aware radio and network technologies. In: Proceedings of the 20th International Symposium on Personal, Indoor and Mobile Radio Communications, Tokyo, 2009. 1–5
- 7 Niu Z S, Zhou S, Zhou S D, et al. Energy efficiency and resource optimized hyper-cellular mobile communication system architecture and its technical challenges. Sci Sin Inform, 2012, 42: 1191–1203 [牛志升, 周盛, 周世东, 等. 能效与资源化的超蜂窝移动通信系统新架构及其技术挑战. 中国科学: 信息科学, 2012, 42: 1191–1203]
- 8 Zhou S, Zhao T, Niu Z S, et al. Software-defined hyper-cellular architecture for green and elastic wireless access. IEEE Commun Mag, 2016, 54: 12–19
- 9 Cai S, Che Y, Duan L, et al. Green 5G heterogeneous networks through dynamic small-cell operation. IEEE J Sel Areas Commun, 2016, 34: 1103–1115
- 10 He Q, Xiao L, Zhong X, et al. Increasing the sum-throughput of cells with a sectorization method for massive MIMO. IEEE Commun Lett, 2014, 18: 1827–1830

- 11 Bi W P, Su X, Xiao L M, et al. On rate region analysis of full-duplex cellular system with inter-user interference cancellation. In: Proceedings of IEEE International Conference on Communication Workshop (ICCW), London, 2015. 1166–1171
- 12 Gao Y, Li Y, Yu H, et al. Energy efficient cooperative sleep control using small cell for wireless networks. *Int J Distrib Sensor Netw*, 2015, 2015: 10–20
- 13 He G N, Zhang S Q, Chen Y, et al. Architecture design and performance evaluation for future green small cell wireless networks. In: Proceedings of IEEE International Conference Communications Workshops (ICC), Budapest, 2013. 1178–1182
- 14 Xu X, He G, Zhang S, et al. On functionality separation for green mobile networks: concept study over LTE. *IEEE Commun Mag*, 2013, 51: 82–90
- 15 Wu Q F, Yang X Y, Liu H, et al. Mobile Guardian: a novel positioning and monitoring system for outdoor special users based on GPS. In: Proceedings of IEEE International Symposium on IT in Medicine and Education, Xiamen, 2008. 596–600
- 16 Luan F Y, Zhang Y, Xiao L M, et al. Fading characteristics of wireless channel on high-speed railway in hilly terrain scenario. *Int J Antenn Propag*, 2013, 2013: 188–192
- 17 Chen L, Li H B, Li Y Z, et al. A learning-based channel model for synergetic transmission technology. *China Commun*, 2015, 12: 83–92
- 18 Chen G F, Zhang Y, Luan F Y, et al. Optimization of AP placement in indoor fingerprint positioning. In: Proceedings of IEEE International Conference on ICT Convergence (ICTC), Jeju, 2013. 98–100
- 19 Chen G, Xiao L, Li J, et al. Dynamic base station selection strategies based on location fingerprinting. In: Proceedings of IEEE 4th Asia-Pacific Conference on Antennas and Propagation (APCAP), Bali Island, 2015. 474–477
- 20 Cai S J, Xiao L M, Wang J, et al. Research on periodical opening of small base-stations for energy conservation. *J Tsinghua Univ (Sci & Technol)*, 2016, 56: 111–116 [蔡世杰, 肖立民, 王京, 等. 以节能为目标的小基站周期性开启研究. *清华大学学报 (自然科学版)*, 2016, 56: 111–116]
- 21 Zhang Y. Wideband Wireless MIMO Channel Measurement, Modeling and Prediction. Dissertation for Ph.D. Degree. Beijing: Tsinghua University, 2010. 138–140 [张炎. 宽带无线 MIMO 信道测量, 建模与预测. 博士学位论文. 北京: 清华大学, 2010. 138–140]
- 22 Wang L, Xu X, Wu Y, et al. Sparse code multiple access-towards massive connectivity and low latency 5G communications. *Telecommun Netw Tech*, 2015, 5: 5
- 23 Lian J, Zhou S D, Zhang X J, et al. Low complexity decoding method for SCMA in uplink random access. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), Washington, 2016. 1–6
- 24 Zhang J, Lu L, Sun Y, et al. PoC of SCMA-based uplink grant-free transmission in UCNC for 5G. *IEEE J Sel Areas Commun*, 2017, in press. doi: 10.1109/JSAC.2017.2687218
- 25 Zhou S, Li Y, Zhao M, et al. Novel techniques to improve downlink multiple access capacity for beyond 3G. *IEEE Commun Mag*, 2005, 43: 61–69
- 26 Zhang S Q, Xu X Q, Lu L, et al. Sparse code multiple access: an energy efficient uplink approach for 5G wireless systems. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), Austin, 2014. 4782–4787
- 27 Wang Y, Zhou S, Xiao L, et al. Sparse Bayesian learning based user detection and channel estimation for SCMA uplink systems. In: Proceedings of IEEE International Conference on Wireless Communications & Signal Processing (WCSP), Nanjing, 2015. 1–5
- 28 Meng X, Wu Y, Chen Y, et al. Low complexity receiver for uplink SCMA system via expectation propagation. In: Proceedings of IEEE Wireless Communications & Networking Conference, San Francisco, 2017
- 29 Cheng M, Wu Y, Li Y, et al. PHY abstraction and system evaluation for SCMA with UL grant-free transmission. In: Proceedings of IEEE Vehicular Technology Conference (VTC Spring), Sydney, 2017

Agile coverage and control for hyper-cellular networks

Wenping BI¹, Xiuqiang XU³, Yuan GAO¹, Guofeng CHEN¹, Yan CHEN³ & Shidong ZHOU^{1,2*}

1. *Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;*

2. *National Laboratory of Information Science and Technology, Tsinghua University, Beijing 100084, China;*

3. *Huawei Technologies, Shanghai 450008, China*

* Corresponding author. E-mail: zhousd@tsinghua.edu.cn

Abstract Hyper-cellular network (HCN) architecture was proposed based on the separation of traffic and control coverage. In this paper, we mainly focus on the decoupling structure for traffic and control coverage. We also provide the decoupling principle with accompanying conditions and examples. With this knowledge, we discuss two key problems in the control process. For channel estimation under partial measurement, based on channel correlation between the spatial and time domain, we propose two channel prediction methods. One method is using traffic base-stations (BSs) to periodically measure channels and make decisions. The other is online and offline learning-based channel prediction. As for the massive connectivity problems in control coverage, sparse coding multiple access is adopted as a grant-free access method to support massive connectivity, low latency, and high reliability for control coverage.

Keywords agile coverage, coverage decoupling, channel estimation, grant-free access, sparse coding multiple access

Wenping BI was born in 1988. He received his B.S. degree from the Department of Electronic Engineering at Tsinghua University in 2012. He is currently working toward a Ph.D. degree in the Department of Electronic Engineering at Tsinghua University. His current research interests include wireless communications, full duplex communications, and interference suppression.

Xiuqiang XU was born in 1984. He received his B.E. degree from Nanjing University, China, in 2007 and his M.E. degree from Nanjing University of Aeronautics and Astronautics, China, in 2010. After graduation, he joined the Green Radio Excellence in Architecture and Technology (GREAT) project at Huawei Technologies Co. Ltd. as a research engineer. His current research interests include non-orthogonal multiple access (NoMA) schemes, scheduling, and radio access network design.

Yan CHEN was born in 1982. She received her B.S. and Ph.D. degrees in 2004 and 2009, respectively, from Chu Kochen Honored College, Zhejiang University, and the Institute of Information and Communication Engineering, Zhejiang University, respectively. Immediately after graduation, she joined Huawei Technologies (Shanghai) Co. Ltd. From 2013 through the present,

she has been the technical leader for Huawei 5G air interface design, focusing mainly on non-orthogonal multiple access schemes and grant-free transmission procedures, as well as system designs for ultra-low latency and ultra-high reliability performance.

Shidong ZHOU was born in 1969. He received B.S. and M.S. degrees from Southeast University, Nanjing, China, in 1991 and 1994, respectively, and a Ph.D. degree from Tsinghua University, Beijing, China, in 1998. He is now a professor with the Department of Electronic Engineering, Tsinghua University. His research interests include mobile communication systems architecture, advanced transmission techniques, wireless channel sounding and modeling, radio resource management, and high energy efficient wireless networks.