



非时齐多元协同通信:虚拟实时服务理论及应用^{*}

陈 巍^{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)

摘 要:提出一种全新通信服务体制——虚拟实时服务,可有效发掘和利用负时间轴上蕴含的大量空闲时、频资源,并利用预先推送较大的时延容忍,人为制造业务趋同性,获取无线多播带来的巨大能效和谱效增益。介绍了虚拟实时服务在认知对象、时间尺度、协同范围 3 个层面的拓展,阐明其能效和谱效增益的代价是计算和存储资源的消耗,从而需要通过计算(含认知)、存储和通信单元的多元协同,以最少计算和存储资源,换取能效和谱效的最大提升。虚拟实时服务有望突破现有通信体制及其理论基础对于资源效率的内在束缚。其两大特质——“非时齐”和“多元协同”的结合,有望形成一种创新的非时齐多元协同通信理论。

关键词:无线通信;虚拟实时服务;多元协同通信;频谱效率;认知无线网络;5G

doi: 10.3969/j.issn.1000-0801.2014.07.008

Non-Homogeneous Multi-Unit Collaborative Communications: A Theory and Application for Virtual Real-Time Service

Chen Wei^{1,2}

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

2. Tsinghua National Lab on Information Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: A novel communication service methodology, virtual real-time service, was proposed to discover and utilize the rich temporal and spectrum resource hidden in the negative time axis. Thanks to the delay tolerance of embedded wireless pushing, it could help to unify the content demand of different users, so as to achieve significant spectral and energy efficiency gains of wireless multicasting. The three extensions of cognition object, time scale, and coordination range, induced by virtual real-time service were described. The consumption of computation and storage resource was a cost to be paid for the gains of the energy and spectral efficiency. Hence, the multi-unit collaboration of computation (including cognition), storage, and communications units was highly desired to minimize the cost of computation and storage resource, while maximize the energy and spectral efficiency gains. The proposed scheme holds the promise of overcoming the stringent restrain on the spectral efficiency, which is caused by the current communication architecture and its underlying theory. The union of its two main characteristics, namely, time non-homogenous transmission and multi-unit collaboration, are expected to form a paradigm-shift theory on time non-homogenous multi-unit collaboration communications.

Key words: wireless communication, virtual real-time service, multi-unit collaborative communication, spectrum efficiency, cognitive wireless network, 5G

^{*} 国家自然科学基金优秀青年科学基金资助项目(No. 61322111),国家重点基础研究发展计划(“973”计划)青年科学家专题基金资助项目(No.2013CB336600),创新人才推进计划基金资助项目(No.2013RA2151),北京市科技新星计划基金资助项目(No.Z121101002512051),教育部新世纪优秀人才计划基金资助项目,清华大学高层次人才建设计划(“221”)人才计划基金资助项目



1 引言

无线通信网络的发展离不开频谱资源。根据 Shannon 公式^[1],带宽和功率是决定通信容量的两种基本资源,而带宽的拓展还能够显著提升能量效率,降低无线通信网络的能耗。然而,适宜无线电通信的带宽是非常有限的。近年来,在无线通信行业整体迅速发展的浪潮中,移动互联网业务激增,大量新的无线系统、标准等不断涌现,导致可用的频谱资源越来越稀缺。为实现无线通信与网络产业的可持续发展而进行的各项研究和开发,已经逐渐感受到了因频谱资源不足而带来的根本制约。频谱稀缺和业务激增之间的矛盾,将是未来一段时期内无线领域研究所面临的主要矛盾,从而成为学术界和工业界所共同关注的热点^[2-4]。

总体来看,克服频谱稀缺的方法有 3 类:第一类是提升无线通信的频谱效率,第二类是避免频谱的闲置与空耗,第三类是从物理学角度发掘新的无线传输频段及复用方式。第一类方法的核心思想是通过先进的物理层技术提升信息传输时单位带宽支撑的吞吐量,从链路层面缓解频谱的稀缺性。其中,物理层的信道编码技术也已逐渐成熟并逼近理论极限^[5,6],近期发展的代表性方向是自适应调制编码^[7]、多天线传输^[8-11]、双向中继网络编码^[12,13]等;第二类方法的核心思想是通过在多个系统间灵活地共享频谱资源,提升频谱的利用率,从系统层面有效缓解频谱的稀缺性,近期关注的代表性方向是认知无线电、动态频谱管理^[14-16]等;第三类方法的核心思想是借助物理学原理,在未开垦的电磁频段上传输信息,或者在已有频段上利用新的物理机理实现多路复用,近期关注的代表性方向是高频段^[17]、太赫兹通信、密集立体覆盖^[18]和极化复用^[19]等。由于第二类方法的动态频谱共享具有实现上的低复杂度和较好的与现有系统兼容的能力,在这一思路下衍生的认知无线电技术成为学术界和工业界所共同关注的热点。

认知无线电能够主动认知、发掘和利用空闲频谱,因而其灵巧频谱使用模式能够有效弥补现有静态频谱分配模式所导致的使用效率低下问题,有效提升无线系统的吞吐量。然而,认知无线电存在一个较为根本的困难,即它难以以为实时业务有效地提供服务质量(QoS)保障。究其原因,是认知无线电只能接入空闲时频资源,而空闲的时、频资源在整个时一空一频维度上是随机的、突发的、零散的、不可控的。当主用户系统的业务繁忙时,其势必占满几乎所

有可用的授权频时。此时,认知用户或二级用户就将面临着在相当长一段时间内找不到空闲频谱可用、无法开展业务传输的尴尬局面。对于有严格时延限制的实时业务来说,将导致服务质量和用户体验的严重下降。为了克服认知无线电的这一缺陷,参考文献[20,21]尝试利用空间维度实现认知无线电对于主用户授权频谱的无害复用,从而在时、频维度被主用户占满的情况下及时传输有最小时延要求的业务。CORE-4 架构^[22]是这一思路的集中体现。但是,伴随着未来无线网络的密集化、立体化趋势,主用户也将逐渐占满空间维度,从而在物理机理的层面导致网络容量的不可扩展性^[23]。

综上所述,维持无线网络的可持续发展,迫切需要在全新的维度上寻找空闲的频谱资源,并且在实际系统中予以有效的发掘和利用。针对这一需求,本文提出了一种在负时间轴上发掘和利用空闲时、频资源的通信体制——虚拟实时服务,并且介绍了其衍生的理论基础——非时齐多元协同通信理论。虚拟实时服务的核心思想是,通过认知、学习和预测用户可能的业务或内容需求,在该用户提出此业务或内容的传输请求之前的“负时间轴”上,利用其中较长时间段内空闲的时、频资源,将用户可能感兴趣的内容提前推送到用户的存储单元中,当用户索取该业务或内容时,只需从存储单元中实时读出即可。注意到,由于内容产生的时刻往往远提前于用户索取的时刻,因此在负时间轴上推送,可以利用的时间尺度显著增加,通过聚合负时间轴上蕴含的大量零碎频谱空洞,可以为认知无线网络带来大量可用的空闲时、频资源。于是,即使认知无线网络只能利用极其零散、突发和不可预测的时频空洞进行非实时的传输,也可以使得用户最终感觉到一种“虚拟实时”的服务。用这种非实时的信息传输支撑用户具有实时感受的服务,就是“虚拟实时服务”这一名称的来源。

本文分析了虚拟实时服务中蕴含的能量和频谱效率以及网络服务质量的提升空间,同时指出了虚拟实时服务所面临的代价,即:认知对象由频谱占用向业务需求的延伸,海量业务由实时传输向提前推送的转变,均导致了认知计算和存储资源的大量消耗。为了逼近计算和存储资源消耗及其与通信资源(带宽、能量)效率之间的最佳折中,虚拟实时服务需要引入通信单元、计算单元(包含认知单元)和存储单元之间的“多元协同”,从而以消耗可扩展的计算和存储资源,实现能效和谐效的大幅提升,节省难以扩展的频谱和能量资源。与此同时,基于空闲时、频资源推

送业务所跨越的时间尺度较大,整个时间跨度内传输环境和资源状态等均可能发生变化,因此传输时间尺度由“时间点”向“时间线”延伸,从而使传输通道具有“非时齐”特性。本文表明,认知无线网络虚拟实时服务体制中的“多元协同”和“非时齐”特质,使得“非时齐多元协同通信”的基础理论及关键技术具有广阔的研究空间和应用价值。

2 虚拟实时服务——通过认知利用负时间轴频谱资源

本节介绍认知无线网络中虚拟实时服务的基本概念、实现方法及其性能优势。在本文的讨论中,通信系统的“时间轴”是展开论述的一个基本出发点,如图 1 所示。在传统通信模式中,系统从用户提出需求的时刻启动业务传输。因此,用户感受到的时延就是从传输启动到完成的时间差。由于实时业务特有的时延约束,其可发掘和利用的空闲时频资源过少,这就会导致业务在规定的时延内无法传输完毕,例如导致移动终端视频播放等业务的中断、卡顿等较差的用户体验,从而影响网络的服务质量。在传统的通信模式中,降低业务时延的主要方法是提升传输的速率,这就需要消耗更多的带宽和能量,这种发展在无线系统的演进中不可持续。

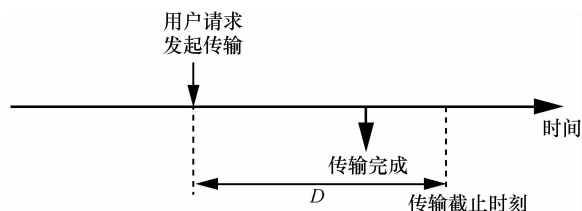


图 1 传统通信体制——始于用户提出传输请求时刻,仅利用正时间轴

2.1 基于单播推送的虚拟实时服务

事实上,在移动互联网的大量应用(如网页浏览、视频播放)中,信源或内容的产生显著超前于用户对于该信源或者内容的索取。为了定量地刻画描述这一现象,将用户提出业务传输请求的时刻定义为整个时间轴的“零点”,如图 2 所示。传统通信理论中,专注于研究和优化正时间轴上的信息传输方法。那么,可以发现信源或内容的产生时刻,实际上位于负时间轴上。不难预见,统计意义上信源的产生往往超前于用户的索取,因此这两个时刻之间具有显著的时间差。在图 2 所示的这一时间差中,往往蕴含了

大量的空闲时、频资源。因此,可在较大时间尺度内充分感知、发掘和聚合空闲时隙。

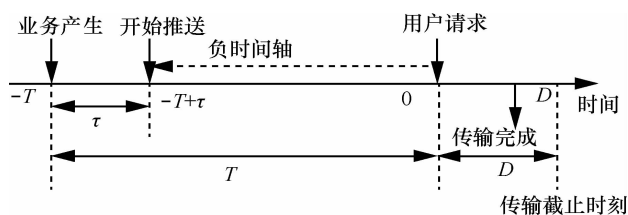


图 2 虚拟实时服务——预先推送用户所需的业务,有效发掘负时间轴的时频资源

若要利用负时间轴上的空闲频谱资源,具体的实现方法就是在负时间轴上,预先将业务“推送”到终端,则用户索取该业务时就可直接从本地存储中读出。特别值得注意的是,空闲时频资源分布具有显著的非均匀性,例如在凌晨时刻用户通信请求少,频谱较空闲。若利用这一时段中的空闲频谱资源进行推送,恰好可以充分利用凌晨的空闲资源,非常适宜推送大量数据至用户终端,从而可以通过非实时的通信来实现虚拟的实时服务。

与此相反的是,经典认知无线电是在用户提出通信请求后才寻找空闲时隙,由于业务 QoS 的时延限制,在短暂的时段内可发掘的空闲时频资源较少。而回到刚才的例子,因凌晨的通信请求较少,故它并不能充分利用此时的空闲频谱。虚拟实施服务的引入,使得认知无线网络的推送无需用户实时发起,因此,“虚拟实时服务”与认知无线电的结合,提供了一条全新的频谱共享思路,借助“虚拟实时服务”,频谱感知能更有效地提升频谱效率,并利用负时间轴传输大量信息。

为了在科学上严格描述这一过程,建立了一个简单的数学模型并实现一些初步的理论分析。主要数学符号如图 2 中所示,符号的具体含义解释如下:

- D ——业务的时延限制;
- T ——用户需求相对于业务产生的时延;
- S ——业务传输所需的时间;
- τ ——计算单元通过预测做出推送判断的耗时;
- p ——发现空闲时隙的概率(对于主用户系统, $p=1$)。

结合图 2 及上述符号,下面初步说明虚拟实时服务的性能提升机理。

时间零点是用户提出业务需求的时刻,因此业务产生时刻为 $-T$,要求的传输截止时刻为 D 。因此,业务必须在时刻 D 之前传输完毕,否则会因信息的过期被认为是“失效



的”,不满足用户的 QoS 需求。由于认知计算单元需要耗时做出推送的决定,因此推送的起始时刻在 $-T+\tau$ 。

接下来分析服务质量得以保障的条件。从推送开始到传输截止时刻,其间的总时长为 $D+T-\tau$ 。注意到发现空闲时隙的概率为 p ,因此,无法在截止时刻之前完成业务传输的概率,等同于在 $D+T-\tau$ 个时隙中空闲时隙少于 S 的概率。于是,业务传输失效的概率为:

$$p_{\text{out}} = \sum_{k=0}^{S-1} \binom{D+T-\tau}{k} p^k (1-p)^{D+T-\tau-k} \quad (1)$$

与此相对的,传统认知无线电的业务传输失效概率为:

$$p_{\text{out}} = \sum_{k=0}^{S-1} \binom{D}{k} p^k (1-p)^{D-k} \quad (2)$$

一般来说, $T-\tau \gg D \geq S$ 。其中,第二个不等号成立的原因:若其不成立,则经典认知无线电无论如何也不能确保在给定的时延要求内完成传输。因此, $D+T-\tau \gg S$ 。此时不难看出 $p_{\text{out}} \approx 0$,对于绝大多数业务,当用户提出传输请求时,都可以直接从本地缓存中提取,从而以非实时的通信,支撑虚拟的实时服务。

对于主用户系统, $p=1$ 。对于此类系统的虚拟实时服务,当且仅当 $D+T-\tau \geq S$ 时,可满足 QoS。而传统通信系统中,当且仅当 $D \geq S$ 时可满足 QoS。显然,虚拟实时服务可以支撑更小的业务演示要求 D ,或者更大的业务量,即更大的传输时间 S 。上述简要的定量分析,就可以直观地展示出虚拟实时服务所带来的资源节省和服务质量提升。

2.2 基于多播推送的虚拟实时服务

在本节中,进一步介绍虚拟实时服务与多播结合,能够带来的显著能效和谱效增益,对第 2.1 节讨论的基于单播推送的虚拟实时服务实现有效拓展。

无线环境的信号传输具有典型的广播特性,因此多播可以利用一组信号,同时向多个终端传输具有相同需求的业务,从而极大地节省了能量消耗和频谱需求。然而,在传统的无线通信网络中,适用于多播传输的业务所占的比重非常有限,因此限制了应用多播来提升能效和谱效的尝试。为何在网页浏览、视频流等应用中,大量用户索取的是相同的业务,而多播的应用范围却非常有限呢?究其原因,根本上是传统无线通信中,若用户在不同时刻索取相同业务,基站需即时发起传输以保证实时性,造成同一业务的反复多次传输。如果想要获得无线环境的多播增益,需适当延迟对用户需求的响应,“对齐”其业务传输。但是,若强

行延迟某些用户的业务传输,以便与其他用户对齐后进行多播,这种在用户提出需求之后的延迟会导致部分用户时延较大,尤其是最先提出业务请求的用户感受到的时延最大,严重损害 QoS。

本文提出,基于“虚拟实时服务”新体制,允许在不影响用户的服务质量和体验的前提下,有效聚合更大时间尺度内的相同业务或内容需求,从而人为制造多播所需的“业务趋同性”,把传统网络中的非多播业务用多播方式进行传输,实现传输次数及能量、频谱资源消耗和占用的“压缩”。

具体的分析如图 3 所示,利用虚拟实时服务解决上述问题的潜力,蕴含在对于负时间轴的利用上。由于信源产生和业务需求的时间差较大,因此在推送时调整业务的传输时刻,实现业务的“时间对齐”。如图 3 所示,当推送的时刻早于用户提出传输请求的时刻时,则这种通过延迟推送时刻“对齐”业务推送的策略,并不会影响用户实际的时延体验。注意到,由于多播的能耗和频谱占用与接收多播业务的终端或用户数量无关,因此,其能量和频谱效率线性正比于接收用户的数量。接收用户的数量不能无限增长,而是受到了图 3 所示的可以“对齐”的时间区间的约束。允许对齐推送的时间区间越长,则多播接收组的用户数量越大。

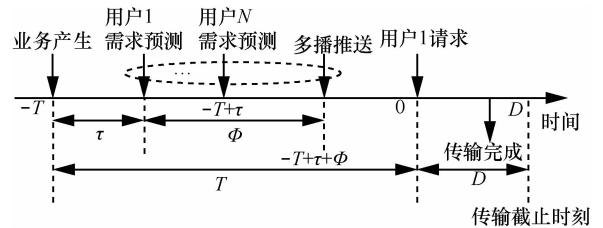


图 3 虚拟实时服务的多播推送——利用负时间轴对齐业务,人为制造“业务趋同性”

为了定量地刻画上述论断,下面简要分析虚拟实时服务体制能带来的多播增益。假设用户对相同业务需求的产生在整个时间轴上是均匀的。为了获取多播带来的能效和谱效增益,需要延迟对于一些用户业务传输请求的响应。定义如下: Φ ——为实现多播而允许的最大响应时延。

换言之,时长 Φ 的时段内的业务将被用同一组信号多播传输完成,该时长正比于多播带来的能量和频谱增益。对于虚拟实时服务,由于需要计入为了实现多播而“人为”引入的时延,多播推送实际允许的时延降低 Φ ,变为 $D+T-$

$\tau - \Phi$ 。类似第 2.1 节的分析,在这一时间段内找到少于 S 个空闲时隙的概率为:

$$p_{\text{out}} = \sum_{k=0}^{S-1} \binom{D+T-\tau-\Phi}{k} p^k (1-p)^{D+T-\tau-\Phi-k} \quad (3)$$

式(3)给出的即 QoS 无法保障的概率。对于传统认知无线电,计入多播带来的允许时延的折损 Φ ,其 QoS 无法保障的概率为:

$$p_{\text{out}} = \sum_{k=0}^{S-1} \binom{D-\Phi}{k} p^k (1-p)^{D-\Phi-k} \quad (4)$$

不难看出,随着允许时延 Φ 的增加,业务失效的概率也随之增大,这就形成了能效、谱效和业务 QoS 之间的矛盾及其折中关系。

考虑一个相对简单的场景,即主用户系统,其 $p=1$ 。此时,对于虚拟实时服务,其 QoS 得以保障的条件变为 $D+T-\tau-\Phi \geq S$ 。因此,为实现多播而允许的最大“人为”业务推送时延为:

$$\Phi_{\text{max}} = D+T-\tau-S \quad (5)$$

而对于传统认知无线电,其 QoS 保障条件为 $D-\Phi \geq S$,其允许的最大“人为”业务推送时延为:

$$\Phi'_{\text{max}} = D-S \quad (6)$$

回顾之前的论述,知道这种人为的推送时延正比于能效和频谱效率增益,因此,虚拟实时服务可以提供的多播增益表示为:

$$\eta = \frac{D+T-\tau-S}{D-S} \quad (7)$$

如第 2.1 节所讨论的, $T-\tau \gg D \geq S$,因此虚拟实时服务可以提供的多播增益的近似下界为:

$$\eta > \frac{T-\tau}{D} \quad (8)$$

由式(8)可见,虚拟实时服务于多播推送的结合,所带来的能效和频谱效率增益是非常惊人的。综上所述,基于业务需求认知和预测的多播推送,构成了一种全新的通信体制——虚拟实时服务。理论分析表明,虚拟实时服务可以极大地发挥认知无线网络的优势,充分地发掘和利用较大时间尺度内的空闲频谱资源,并且最大地发挥多播所能带来的能效和谱效增益。从理论创新角度来看,负时间轴概念的提出,不仅可以量化描述虚拟实时服务的内在机理,同时还可以对经典的无线 QoS 理论^[24]进行有效的推广和扩充,从而量化分析虚拟实时服务的 QoS 性能,限于本文篇幅,这里不展开赘述。

3 非时齐多元协同通信——三项拓展形成一个理论

第 2 节从定性和定量两个角度对虚拟实时服务这种创新通信体制进行了阐述和分析,从中可以看出,虚拟实时服务蕴含着显著提升能效和谱效的巨大潜力。然而,对于该体制所带来的额外代价,尚未进行有效的分析和量化。

在本节中,重点从定性和定量角度讨论虚拟实时服务的额外代价。简言之,这种新的通信体制需要消耗更多的计算和存储资源:

- 一方面,终端接收的推送业务须保留在存储单元内,直到用户阅读后,或有效性过期后删除。推送业务越多、存储时间越长,消耗的存储量越大;
- 另一方面,采集、分析大量用户数据,预测其业务需求的认知过程,将产生大量认知计算量,而终端对多播信号,特别是多个频带上多播信号进行解调、解码,将产生接收计算量。

因此,当基站进行海量业务认知,并在较宽的频带上推送大量业务时,存储和计算资源大量消耗的问题就突显出来,接踵而至。虚拟实时服务中,需要有效的协同通信单元、计算单元和存储单元,有效降低对于计算和存储资源的消耗。因此,虚拟实时服务中所需要解决的一个核心问题就是:如何通过计算、存储与通信单元的多元协同,最小化为提升通信资源效率而付出计算、存储的代价?

为了有效阐明上述问题的答案,需要从原理上阐明虚拟实时服务体制相对于传统通信体制的区别和新意。本文将其总结为认知对象、时间尺度、协同范围 3 个层面的拓展,下面分别对其进行介绍,并说明由这 3 个方面的拓展所导致的计算和存储资源消耗内在机理与协同关系。

3.1 认知对象拓展——从频谱占用到业务需求

虚拟实时服务是建立在对于用户的业务需求预测的基础上的,这就导致了认知对象的拓展,并因此需要付出更多的计算资源。传统认知无线电,认知的对象是单纯的频谱占用情况。而认知在虚拟实时服务中则需要认知用户的业务偏好,发掘用户群体的兴趣趋同特性,准确预测需求,通过发掘用户需求进行预测,从而推送其未来最有可能索取的业务。因此,认知的对象将由频谱占用拓展到用户需求、群体偏好等。



虚拟实时服务可准确而迅速地预测用户的业务需求,包括内容对象和需求时间。对于需求对象的预测错误将导致基站推送用户不需要的业务,从而浪费通信资源,包括带宽和能量;对于用户需求时间的预测错误将会导致推送过早而空耗存储资源,或者推送过晚而错失 QoS 所需的截止时间。以第 2.1 节所述的单播推送为例,若能够较为准确地预测用户对用户需求相对于业务产生的时延 T ,则可以最优化推送的时机,使得推送完成的时刻恰好提前于用户提出需求的时刻,从而减少对存储介质的消耗。若预测错误,推送了用户根本不需要的内容,则相当于 $T \rightarrow \infty$,这将导致存储介质的高度浪费,而所消耗的通信资源(能量、带宽)均被浪费。

如第 2.1 节所示,计算单元认知和预测用户业务需求的耗时是 τ ,这个参数是协同计算单元和通信单元的桥梁。显然,耗时 τ 越长,则留给推送的可用时间资源 $D+T-\tau$ 越少,因此通信系统的性能会变差。但是,在给定的计算资源下,认知计算的耗时越长,则预测得更为准确,可以有效避免推送用户根本不需要的内容,空耗通信和存储资源。因此,虚拟实时服务需要给定一个业务预测的准确率,在这一约束下,认知计算时延 τ 和投入的计算资源(例如 CPU 数量、线程 ξ 等)成反比,即:

$$\tau \propto \frac{1}{\xi} \quad (9)$$

式(9)对于计算资源共享的无线接入网络,如近年来受到关注的 Cloud-Ran,具有参考价值。当计算资源共享时,每个用户可用的计算资源,反比于系统中共享计算资源的用户数量 N ,因此虚拟实时服务的计算时延正比于享受推送服务的用户数,即:

$$\tau \propto N \quad (10)$$

虚拟实时服务中,认知对象从频谱占用到业务需求的拓展,使得通信理论与机器学习、人工智能和社交网络等进一步交叉融合,也产生了一些跨学科的开放问题。例如:对于用户业务需求时刻的预测精度和总认知计算量投入之间的关系,就需要引入机器学习等方面的方法,以便阐明其折中机理和性能极限。

3.2 时间尺度拓展——从时间“点”到时间“线”

由于虚拟实时服务的信息传输过程始于负时间轴上的时间点,因此这种通信体制可以在非常大的时间尺度上聚合空闲频谱资源,并且通过对齐传输位置获取多播增益。这种时间尺度上的拓展,就使得通信过程实现从时间

“点”到时间“线”的延伸,并且使得存储单元及其资源消耗成为一个不可忽略的因素。

回顾经典信息论,其中传输环境、资源状态保持恒定的时间尺度要远大于信息分组、编码长度等的时间尺度,因此经典通信理论考虑的是“时间点”。在虚拟实时服务中,由于利用了负时间轴而在更大的时间尺度上聚合资源,基于空闲时、频资源推送业务所跨越的时间尺度较大,整个时间跨度内传输环境和资源状态等均可能发生变化,因此信息分组、编码长度等信息传输相关的时间尺度远大于传输环境、资源状态保持恒定的时间尺度。故此,可以认为虚拟实时服务的通信过程需要在时间“线”上进行描述,并将传输理论从“时间点”拓展至“时间线”,以便有效刻画其传输环境和资源状态变化所产生的“非时齐”特性。

在时间线上的传输过程中,经典信息论的恒参数统计信道不再适用。下面用一个简单的例子来初步展示其差异。回顾传统的对称二进制信道(BSC),其差错概率为 ε ,同时收发双方均知晓 ε 的值,从而能够自适应调整速率和编码方式。这一模型难以适用于虚拟实时服务。具体来说,虚拟实时服务于漫长的时间线上,选取分散度较高的空闲时隙进行通信,可能在每几个信道使用(channel use)之间具有较大的时间间隔,从而导致这些信道使用的状态和差错概率有很大的差异。换言之,同一个信道编码的码块的各个比特,将经历不同差错概率的 BSC,其差错概率将是序列号的函数 ε_k ,且 ε_k 未必具有某种统计特性,这种“非时齐”信道如图 4 所示。

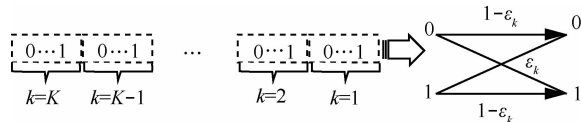


图 4 “非时齐”对称二进制信道

需要指出的是,经典信息论中对于时变信道的处理方法不能适用于如图 4 所示的非时齐信道。在经典信息论中,对于时变信道的处理方法是将其分割为多个不同信道状态的子信道。针对一个状态固定的子信道,可以假设其中传输的符号足够多,以至于统计规律成立,因此给定信道状态条件后的容量可以用典型集为基础进行分析计算,之后再对不同状态求平均即可,其思路植根于将时间线拆分成时间点进行分析。然而,对于如图 4 所示的非时齐信道,其每个状态下所传输的符号极少,因此难以用典型机

来估算其瞬时容量,因此时间线无法拆分成时间点,必须探索适应于此类时间线传输过程的全新通信理论。已有研究证明,非恒速编码(rateless coding)是一类有效应对信道非时齐特性的编码方式。

虚拟实时服务中,通信的时间线拓展的另一个重要影响是导致了存储单元及其资源消耗在系统中直接上升为主要矛盾。传统通信模式中,业务或码块的传输时间短,且传输完成后即完成内容的使用而删除,因此不会占用过多的存储资源。当通信过程拓展至时间线上后,所传输的码块需要长时间存储直至可以解码,而信息内容直至用户读出使用后会删除。因此,必然会消耗较大的存储资源,以之换取通信中能量和频谱效率的提升。

接下来,对存储资源的消耗做简要的定量分析。为分析便利,近似认为存储资源的消耗从业务推送完成后开始,至用户读出后终止。换言之,存储资源的消耗量不仅与推送的业务量有关系,也决定于推送的业务在缓存中存储的时间。考虑一种简单的情况,基站每隔时间 Δ 向一个用户推送一个分组,每个分组包含 P 个比特,推送完成到用户读取的时间差为 θ ,则对于一个用户终端所需的存储空间大小为:

$$\Xi = P \times \left\lceil \frac{\theta}{\Delta} \right\rceil \quad (11)$$

显然,内容推送得越频繁,每次推送的业务量越大,且推送的提前量越大,则终端所消耗的存储资源量也就越大。对于主用户系统($p=1$),式(11)中推送完成到用户读取的时间差 θ 可以由进一步的定量分析得到。由第2.1节可知:

$$\theta = \max\{T - \tau - S, 0\} \quad (12)$$

从中不难看出,通信服务质量及频谱效率与存储资源的消耗之间存在着某种折中关系,降低业务传输超时概率,往往会导致存储资源的更多消耗。换言之,虚拟实时服务所带来的频谱效率提升,是以存储资源的消耗为代价的。此外,当虚拟实时服务引入中继节点时,由于各跳信道在时间线上的容量不同步,因此中继的瞬时存储量是一个有涨落的随机变量,甚至可能导致信息在中继节点的大量积压,从而对中继节点的存储资源提出严苛的挑战。

3.3 协同范围拓展——从多域协同到多元协同

回顾第3.1节和第3.2节不难看出,虚拟实时服务通过引入认知计算和存储单元的有机协同,以消耗可扩展的计算和存储资源为代价,换取严格受限的能量和频谱效率的提升,实现认知无线网络的可持续发展。上述分

析中还表明,计算资源、存储资源和通信资源之间,存在着三角形的折中关系,一方的优化要以另一方的额外付出为代价。因此,3种资源的最佳折中是什么?即为了达到通信资源的效率最优,所消耗的计算和存储资源最少是什么?上述问题就构成了虚拟实时服务的理论性能极限的刻画。而若要达到这种最佳折中,认知计算、存储和通信这3个单元就需要相互协同,从而实现资源的最优配置。

将认知计算、存储和通信这3个单元的协同称为多元协同。多元协同是传统无线网络中协同思想的进一步拓展。回顾一下,传统的无线网络协同主要体现在如下几个方面。

- 多层协同:协同在物理层、链路层、网络层、传输层乃至应用层以及天线、处理电路等功能层,或模块域之间展开,旨在提升通信网络的整体性能。
- 多用户协同:协同在不同用户、终端、设备之间展开,旨在通过相互信息转发、联合处理等获取空间分集增益和阵列增益,提升链路的有效性和可靠性。
- 多网协同:协同在不同的无线网络、体制和标准之间展开,旨在通过有效的资源管理抑制干扰、共享频谱,获得空间频谱复用增益,提升单位面积的频谱效率。

与上述3类常见的协同方式相比,多元协同主要在通信系统内部,如功能层、节点、各类资源之间展开,因而与之相关的性能参数、运行机理和资源消耗都在信息论、网络理论、资源管理理论中有较成熟的描述。

本文所提出的虚拟实时服务这一全新体系构架中,势必要以消耗大量计算(含认知)和存储资源,使得这些资源成为制约系统性能的瓶颈,第3.1节和第3.2节的分析也有力地说明了这一点。因此,协同的范围从通信系统内部的多域,延拓到了通信系统之外的计算、存储单元等多元。此问题是经典信息与通信理论中从未涉及的全新方向,也不同于异构无线网络中已有的协同问题。在传统通信理论中,并没有将系统的性能与计算和存储单元的资源消耗建立起深入的内在联系,较少考虑上述通信、计算、存储单元之间的跨单元协同,因此在传统的通信与网络理论基础上需要进一步地延拓和创新。图5对多元协同的重要折中关系进行了简要的归纳。从图5中和第3.1节、第3.2节分析均可看出,消耗计算和存储资源可换取能量和频谱等通信资源效率的提升;而计算能力的提升,对业务需求预测的精准,还能带来存储资源的节省,并避免通信资源的浪费。因此,学习、认知技术在未来无线通信系统中具有重要作用。

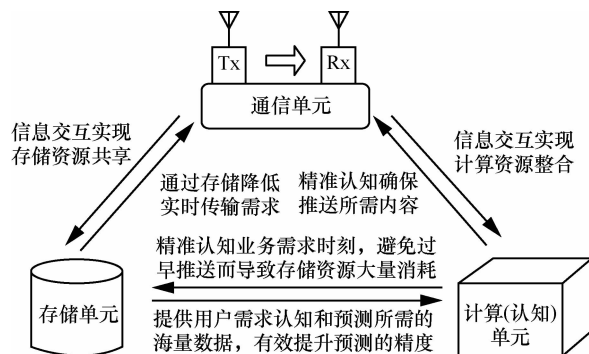


图5 通信单元、计算(认知)单元、存储单元的协同与资源消耗的折中关系

多元协同中需解决的核心问题是:在给定的业务感知精度下,阐明频谱效率、能量效率与计算、存储资源消耗之间的最优折中及其逼近方法。除此之外,第3.2节中所提到的“非时齐”特性,对于通信单元的信息传输理论具有深刻的影响和改变,将“非时齐”和“多元协同”结合起来,就形成了一种适用于虚拟实时服务认知无线网络的新研究方向——非时齐多元协同通信理论。

4 结束语

随着移动互联网的迅速发展和普及,移动业务的海量化、多样化以及移动用户的密集化、立体化,正在导致无线移动通信面临着日益突出的频谱稀缺、能耗居高不下问题。回溯信息论的物理学基础可以看到,进一步拓展无线频谱利用效率和复用维度是未来无线通信发展的必经之路。然而,单纯依靠发掘用户业务请求之后的正时间轴上的空闲频谱资源,或提升其频谱效率,所能带来的容量增益已经捉襟见肘,难以支持未来移动互联网的可持续发展。

本文提出了一种全新的通信服务体制——虚拟实时服务,通过有效预测用户的业务需求,可在用户提出业务传输请求之前的负时间轴上,预先将业务推送到用户设备的缓存区域,从而以非实时的通信,使用户感受到虚拟实时的服务。虚拟实时服务能够有效发掘和利用分布于负时间轴上的空闲时、频资源,特别是充分利用凌晨等空闲时段所蕴含的大量频谱资源。同时,新体制还能够在不影响虚拟实时性的前提下,通过对齐业务的推送传输时间,人为制造“业务趋同性”,将传统上只能单播传输的业务合并进行多播推送,从而借助多播增益大幅度提升能量和频谱效率。

本文初步从定性、定量两个角度分析了虚拟实时服务中,所能带来的通信系统的显著性能增益,并阐明了虚拟

实时服务在认知对象从频谱占用到业务需求、时间尺度从时间点到时间线,协同范围从多域协同到多元协同3个方面的深刻拓展。本文阐明了为获得通信性能增益,虚拟实时服务所付出的代价是消耗大量的计算(认知)和存储资源,换言之,它尝试了一种无线通信的可持续发展模式——以可扩展的计算和存储资源换取难以扩展的能量和频谱资源。为了以最小的计算和存储资源代价,换取能量和频谱效率的最大化提升,虚拟实时服务需要有效开展计算(认知)单元、存储单元和通信单元的多元协同。在信息传输从时间点到时间线的延伸中还发现,信息编码时间尺度大于信道和网络参数保持恒定的时间,出现了经典信息论较少谈及的非时齐信息传输。为了应对虚拟实时服务3方面拓展所带来的挑战,本文综合其中“非时齐”和“多元协同”特色,提出了一种全新的研究框架——非时齐多元协同通信理论。

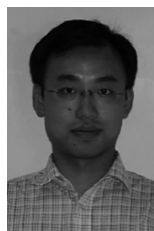
虚拟实时服务体制及依托其衍生的非时齐多元协同通信理论,从全新的维度给出了认知无线网络的频谱和能量效率的进一步提升空间,对其QoS保障给出了有效方法。同时,相关探索启发了无线通信与信息理论的资源拓展和多元协同时代,从基础理论、系统构架和关键算法3个层面,推进未来移动互联网的全面演进。

参考文献

- Shannon C. Mathematical Theory of Communication. Illinois: University Illinois Press, 1963
- Tse D, Viswanath P. Fundamentals of Wireless Communication. Cambridge: Cambridge University Press, 2005
- Gupta P, Kumar P R. The capacity of wireless networks. IEEE Transactions on Information Theory, 2000, 46(2): 388~404
- Akyildiz I F, Wang X, Wang W. Wireless mesh networks: a survey. Computer Networks, 2005, 47(4): 445~487
- Richardson T J, Shokrollahi M A, Urbanke R L. Design of capacity-approaching irregular low-density parity-check codes. IEEE Transactions on Information Theory, 2001, 47(2): 619~637
- Berrou C, Glavieux A. Near optimum error correcting coding and decoding: Turbo-codes. IEEE Transactions on Communications, 1996, 44(10): 1261~1271
- Goldsmith A J, Chua S G. Variable-rate variable-power MQAM for fading channels. IEEE Transactions on Communications, 1997, 45(10): 1218~1230
- Foschini G J, Gans M J. On limits of wireless communications in a fading environment when using multiple antennas. Wireless Personal Communications, 1998, 6(3): 311~335

- 9 Alamouti S. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 1998, 16(8): 1451~1458
- 10 Karakayali M K, Foschini G J, Valenzuela R A. Network coordination for spectrally efficient communications in cellular systems. *Wireless Communications, IEEE*, 2006, 13 (4): 56~61
- 11 Cadambe V R, Jafar S A. Interference alignment and degrees of freedom of the-user interference channel. *IEEE Transactions on Information Theory*, 2008, 54(8): 3425~3441
- 12 Chen W, Letaief K B, Cao Z. A cross layer method for interference cancellation and network coding in wireless networks. *Proceedings of IEEE International Conference on Communications*, 2006, ICC'06, IEEE, 2006: 3693~3698
- 13 Katti S, Rahul H, Hu W, *et al.* XORs in the air: practical wireless network coding. *IEEE/ACM Transactions on Networking (TON)*, 2008, 16(3): 497~510
- 14 Akyildiz I F, Lee W Y, Vuran M C, *et al.* Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks*, 2006, 50(13): 2127~2159
- 15 Zhao Q, Sadler B M. A survey of dynamic spectrum access. *Signal Processing Magazine, IEEE*, 2007, 24(3): 79~89
- 16 Haykin S. Cognitive radio: brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 2005, 23(2): 201~220
- 17 Gutierrez F, Agarwal S, Parrish K, *et al.* On-chip integrated antenna structures in CMOS for 60 GHz WPAN systems. *IEEE Journal on Selected Areas in Communications*, 2009, 27(8): 1367~1378
- 18 陈巍,王东明,王家恒等.密集立体覆盖:未来移动通信的机遇和挑战. *电信科学*, 2013, 29 (6): 2~9
- 19 Andrews M R, Mitra P P. Tripling the capacity of wireless communications using electromagnetic polarization. *Nature*, 2001, 409(6818): 316~318
- 20 Liu J, Chen W, Cao Z, *et al.* Cooperative beamforming for cognitive radio networks: a cross-layer design. *IEEE Transactions on Communications*, 2012, 60(5): 1420~1431
- 21 Wang X, Chen W, Cao Z. ARCOR: agile rateless coded relaying for cognitive radios. *IEEE Transactions on Vehicular Technology*, 2011, 60(6): 2777~2789
- 22 Wang X, Liu J, Chen W, *et al.* CORE-4: cognition oriented relaying exploiting 4-D spectrum holes. *Wireless Communications and Mobile Computing Conference (IWCMC)*, 2011 7th International, IEEE, Turkey, Istanbul, 2011: 1982~1987
- 23 Franceschetti M, Migliore D, Minero P. The capacity of wireless networks: information-theoretic and physical limits. *IEEE Transactions on Information Theory*, 2009, 55(8): 3413~3424
- 24 Hou I H, Borkar V, Kumar P R. A theory of QoS for wireless. *Proceedings of INFOCOM 2009, IEEE, Brazil, Rio de Janeiro*, 2009

[作者简介]



陈巍,男,博士,清华大学电子工程系教授、博士生导师、系教学委员会主任、系务会成员,主要研究方向为无线通信、信息论、移动互联网。

(收稿日期:2014-05-20)

项目简介

陈巍,男,清华大学电子工程系教授、博士生导师、系教学委员会主任、系务会成员,清华大学深圳研究院主任研究员(兼职),北方工业大学客座教授(兼职),中国通信学会青年工作委员会副主任委员,中国电子学会教育工作委员会青年工作组副主任委员,北京市通信学会青年工作委员会委员,IEEE Transactions on Education, IEEE Wireless Communications Letters, China Communications 等国际期刊编委,“全国五一劳动奖章”、“北京市优秀教师”荣誉称号获得者,国家重点基础研究发展计划(“973”计划)首席科学家,第十四届霍英东高校青年教师奖一等奖、国家自然科学基金优秀青年科学基金获得者,IEEE 高级会员,国家

中青年科技创新领军人才计划、教育部新世纪优秀人才支持计划、北京市科技新星计划和清华大学“221”人才计划入选者。

陈巍教授于2002年和2007年在清华大学分别获得学士和博士学位(优秀毕业生、优秀学位论文)。2005-2007年,作为访问学生在香港科技大学进行联合培养。2007年7月留校任教,2010年晋升副教授,2012年31岁时破格晋升教授。任教后于2010年6-9月受UK-China Science Bridge Fellowship资助访问英国南安普敦大学,2014年6-9月受法国政府资助访问巴黎高等科技工程师学校集团(Institute Mines-Telecom),此外多次访问香港科技大学、香港



中文大学等。目前主要从事无线通信、信息论与网络理论等方面的研究工作。曾任 IEEE VTC-S 2011 技术委员会主席、IEEE CCNC 2013 无线分会技术委员会主席、IEEE ICC 2010 无线分会技术委员会主席、IEEE ICC 2014 通信前沿分会技术委员会主席、IEEE COCONET 2008 技术委员会主席、IEEE ICC 2013 Tutorial Chair (导论课程主席)、IEEE ICC 2012 Publication Chair (出版事务主席)、IEEE ICC 2008 Student Travel Grant Chair (学生旅费资助评选委员会主席)。

因在协同通信与网络信息论方面的学术贡献,荣获第五届 IEEE 通信学会亚太区最杰出青年学者奖,因在协同通信资源管理的公平性方面的学术贡献,获 2009 年 IEEE 马可尼论文奖,暨 IEEE Transactions on Wireless Communications 年度最佳论文奖,这是该奖项自设立以来首次授予在北美以外工作的学者。此外还曾获得 2012 年 IEEE 智能电网通信会议最佳论文奖,2007 年 IEEE 国际跨层设计会议最佳论文奖,2006 年 IEEE 国际通信大会最佳论文奖等 IEEE 会议最佳论文奖。2011 年获清华大学学术新人奖(清华大学设立的 40 岁以下青年教师最高学术荣誉)。在教学方面,曾获首届全国高校青年教师教学竞赛工科组一等奖第一名,第七届北京市青年教师教学基本功比赛一等奖、最佳教案奖,2012 年清华大学青年教师教学优秀奖,第四届清华大学青年教师教学大赛一等奖等荣誉。

随着移动互联网时代的来临,无线通信的发展遇到了业务量激增、频谱枯竭、能耗居高不下等困境,而传统点到点传输已经逼近其理论性能极限,通过增加蜂窝密度提升谱效的方法也开始受到站址资源、干扰、基站附加基础功耗和切换等诸多问题的挑战。无线网络在目前的能效和谱效下,不仅难以容纳快速增长的新业务,还逐渐成为高能耗的信息基础设施。因此,亟需在点到点传输逼近 Shannon 界的情况下,寻求进一步大幅提升频谱和能量效率的全新思路,支持无线通信的可持续发展。

国家自然科学基金优秀青年科学基金项目“认知无线网络”提出了在“负时间轴”上发掘空闲频谱资源的研究思路,从全新的频谱利用维度上提升无线通信的频谱使用效率,并以此换取能量效

率的同步提升。本项目提出,认知无线网络的认知对象应当实现由频谱占用到用户需求的拓展,通过认知、学习和预测用户可能的业务或内容需求,就可以利用凌晨等频谱大量空闲的时段,在用户提出业务请求之前的“负时间轴”上,预先将用户可能感兴趣的内容提前推送到用户的存储单元中,当用户索取该业务或内容时,只需从存储单元中实时读出即可。

认知推送可将认知无线网络局限于“时间点”上的传输拓展到“时间线”上,延伸了发掘利用空闲频谱资源的时间尺度,避免了传统认知无线网络在用户发起请求后,才在“正时间轴”上发掘空闲频谱的仓促局面和由此引发的业务时延,有效地以非实时的信息传输支撑用户具有实时感受的服务——即“虚拟实时服务”。此外,对于“负时间轴”的利用,还巧妙地实现了一种人为制造“业务趋同性”的可能——在更大时间尺度内调整对齐对不同用户的推送时机,在不影响用户的服务质量和体验的前提下,把传统网络中的非多播业务用多播方式进行传输,实现传输次数及能量、频谱资源消耗和占用的“压缩”。

本项目注意到:认知无线网络的虚拟实时服务体制所带来的能量和频谱效率提升,是以认知计算和存储资源的消耗为代价换取的,因此着力研究计算、存储和通信单元三者之间的最佳协同方式,将协同的范围、规模由通信网内的多域协同拓展到计算、存储、通信系统之间的多元协同,追求以最少的计算、存储资源消耗,换取能量和频谱效率的最大提升。

本项目依托于清华大学微波与数字通信技术国家重点实验室/信息科学与技术国家实验室(筹)开展研究工作,力争通过为期 3 年的工作形成统一的“非时齐多元协同通信理论”。实验室拥有码流仪、超高速示波器、逻辑分析仪、宽带信道仿真仪、宽带矢量网络分析仪等一大批高端测试仪表设备。2011 年底,清华大学电子工程馆(1 万余平方米)落成并投入使用,为学术团队提供了良好的办公、实验环境。目前,整个团队已拥有超过 160 m² 的实验室面积,形成了一只包括 2 名教授、1 名讲师以及多名博士后、博士生和硕士生在内的层次合理的科研团队。