

无线网络中一种基于粒子化接入思想的 细粒度延迟保障机制研究

孙珂, 李波, 闫中江, 杨懋

(西北工业大学, 电子信息学院, 陕西 西安 710072)

摘 要:针对目前无线网络中时延保障机制存在的若干问题(即可扩展性差、服务等级粒度粗、以牺牲资源利用率的代价来换取时延性能提升等),提出了粒子化接入的思想和相应的接入机制。将业务流建模为一群携带了一定信息量并在一定时间内有效的信息粒子,给出了信息粒子和信息粒子群的定义。给出了满足信息粒子群的时延要求所需的最小可达接入带宽的计算方法。证明了使用 EDF(earliest deadline first)策略可以达到最小可达接入带宽。提出了一种在无线网络接入动态环境下,基于粒子化思想的细粒度延迟保障机制。针对无线网络下行应用场景进行了仿真性能验证。仿真结果表明,在高业务负载情况下,相较于传统接入机制中平均时延和丢包率的快速增长,粒子化接入策略依然能稳定提供较优的时延保障和有效吞吐量。研究成果为后续进一步完善无线网络的服务质量保障机制提供了新的思路。

关 键 词:无线网络;接入机制;粒子化接入;时延保障

中图分类号: TN925

文献标志码: A

文章编号: 1000-2758(2023)01-0081-09

近年来,无线网络的发展方兴未艾,蜂窝移动通信网络、Wi-Fi、蓝牙等网络类型都快速发展。同时,人们对于网络所能提供的服务类型需求也日趋差异化和精细化。特别是各种实时业务,对无线网络所能提供的服务质量提出了更高的要求,因此,对如何在无线接入网络中解决确定性时延保障问题的研究也越来越得到人们的重视。

目前的无线网络中时延问题的研究按照所关注的业务类型大致可以分为 2 类,分别为只考虑了单业务类型的研究工作^[1-3]和考虑了多业务类型的研究工作^[4-12]。而在研究多业务类型的方案中还可进一步细分为两大类工作,即“基于流”的机制^[4-7]和“基于类”的机制^[8-12]。

只考虑单业务类型的研究主要通过改善系统总吞吐量、调度策略、排队策略等方式来提升对时延敏

感类业务的服务质量,但这些策略并不能提供确定性的时延保障。文献[1]中的方案,只针对视频流业务。文献[2]中的方案能在减少整体网络延迟和提高速率之间提供良好的折衷。文献[3]通过缩小聚合级别的候选范围,可以更快地搜索到最优聚合级别来改善时延。在目前复杂的网络需求中,这些只考虑了单类型业务研究工作很难满足现实网络中各类业务的不同需求,因此具有一定的局限性。

为了满足现有网络各种业务个性化的需求,近年来有不少研究工作考虑了多业务共存的情况,这些研究中有 2 类主要的解决方法,“基于流”的机制和“基于类”的机制。其中,“基于流”的机制是一种面向业务流提供资源预约服务的服务质量保障机制,其核心思想是在连接业务流源端和目的端的端到端路径所经过的所有转发节点中预留一定量的资

收稿日期:2022-05-20

基金项目:国家自然科学基金面上项目(61771390,61771392,61871322)与航空科学基金(201955053002,20185553035)资助

作者简介:孙珂(1998—),西北工业大学硕士研究生,主要从事无线通信网络基础理论与应用技术研究。

通信作者:李波(1971—),西北工业大学教授,主要从事无线通信网络基础理论与应用技术研究。

e-mail:libo.npu@nwpu.edu.cn

源,以满足相应业务流的资源需求,这种机制具有较好的服务质量性能。文献[4-5]中都使用了 TDMA 的方式来保障时延,为相应的优先级业务流节点预留了资源。其中为了减少预留资源的浪费,文献[4]提出在轮询前先进行一轮 RTS 发送,来确定相应的节点是否有数据需要发送。文献[5]则使用了将 TDMA 的协调层和底层业务分离的思想,将非实时业务穿插在实时业务流的间隙进行发送。另外,文献[6-7]使用预约的方式来保障业务流的传输。文献[6]提出了一种多步预约的方式,提高预约信息的可靠性。为了处理预测不准确的情况,文献[7]中提出灵活预约的思想来调整预约的资源。毫无疑问,“基于流”机制是对“电路交换”机制的一种模拟,能有效保障相应的业务流服务质量。然而,使用“基于流”机制的问题在于预约的资源容易被浪费,且需要网络接入控制节点为相应的每个业务流维持状态信息,导致了可扩展性差的问题。

多业务类型研究中的另一类,即“基于类”的机制,可以在一定程度上解决“基于流”的机制所存在的资源浪费和可扩展性差的问题。“基于类”的机制是一种面向不同分组类型提供差异化服务的服务质量保障机制,其基本思想是将网络中的业务流分组划分为若干类,并在网络的中转节点中为不同类型的分组提供差异化服务。文献[8-11]是通过对相应的业务流赋予不同的优先级来改善实时业务的时延性能。文献[8]明确将业务分为 3 类,最高的优先级可以无需退避直接抢占信道,次一级的有较少的退避等待时间,来保证其优先于最低优先级。文献[9]中设计了一种定价机制来保证所有的网关如实报告实际优先级。文献[10]设定了周期性接入节点优先级(周期越长的优先级越高),提出了一种新的半持久调度算法,以保证满足周期性实时通信的延迟要求,并使用最少的频谱资源。文献[11]也使用了赋予优先级的方式。另外,文献[12]中使用了软件定义网络,可以根据队列的可用带宽动态地为业务流分配队列。但由于“基于类”的机制仅将业务流分组划分为为数不多的几类,故存在着服务质量保障粒度过大问题。而且随着目前不同网络业务需求种类的不断增加,现有的类型划分会无法满足不同业务个性化的服务质量保障。若需要设计和添加新的业务类型,将势必更难平衡与原有各类型之间的“公平性”,这势必导致了其可扩展性差的问题。总之,“基于类”机制的问题在于服务质量保障

的粒度大且可扩展性仍然不高。

针对目前时延保障机制所存在的主要问题,本文提出了粒子化接入的思想和方法,将业务流建模为信息粒子群。并且给出了满足给定信息粒子群中所有信息粒子时延需求的最小可达接入带宽的计算方式以及相应的接入传输方法。基于粒子化思想,提出了一种在无线网络接入动态环境下的细粒度延迟保障机制。该接入机制可以在信息粒子的细粒度上,保障各业务流差异化的时延性能需求,且可同时做到对频谱资源最为充分的利用。

1 粒子化接入基本思想

如图 1 所示,本节从基础模型到接入算法描述了粒子化接入的基本思想。首先说明了基于信息粒子的业务流模型,然后说明了所提出的业务流模型静态情况下最小可达接入带宽和相应的接入传输方法,最后说明了实际情况下业务流动态到达时所采用的粒子化接入时延保障机制的基本思想。

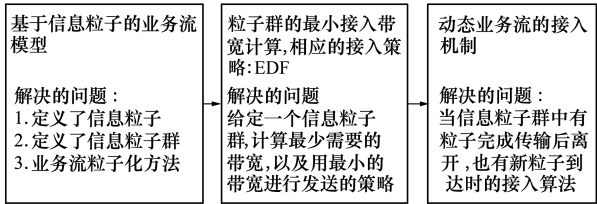


图 1 粒子化接入基本思想示意图

1) 基于信息粒子的业务流模型

保障网络的端到端业务流服务质量的 4 个主要性能指标是:带宽、时延、时延抖动、分组丢弃率。其中,时延指标尤为关键^[13]。因此,本文所提出的基于信息粒子的业务流模型,将围绕时延指标展开。所提出的业务流模型,将聚焦于粒度更细的级别。模型将基于以下几点假设:

- (1) 一个端到端业务流可以被划分为若干个具有一定大小(即信息承载量 I , 单位为 bit) 的信息块;
- (2) 每个信息块均明确对应着属于自己的端到端时延容忍值(即该信息块截止时刻 T , 单位为 s);
- (3) 对于一个承载量为 I 且截止时刻为 T 的信息块,如果网络能在 T 时刻前将其所承载的所有信息比特成功地传送至接收端,则认为对该信息块的传输是有效的。

本文中,将具有承载量属性和截止时刻属性的信息块称为一个信息粒子,而端到端业务流便被建模为由信息粒子构成的信息粒子流。

2) 信息粒子群的最小可达接入带宽的计算及相应的接入传输策略

在粒子化接入机制中,网络中的路由节点需要处理由大量信息粒子所构成的信息粒子群的接入问题。针对这个问题,核心思想是在给定信息粒子群(其包含的所有粒子的属性均为已知)的情况下,找到可以使其中所有信息粒子传输都为有效的最小可达接入带宽。在后续的分析中发现至少存在 EDF 传输策略可在最小可达接入带宽的约束下,实现对所有信息粒子的有效传输。

3) 动态业务流的接入机制

由于在实际应用场景下,每个节点会将队列中的信息粒子不断地发送出去,并且不断有新的信息粒子到达该节点。因此,本文提出的接入机制中,将在每个信息粒子到达发送队列的时候,计算相应的信息粒子群最小可达接入带宽。如果该计算所得到的最小可达接入带宽大于给定的信道带宽,则说明当前新到达的信息粒子将无法被有效发送,反之则将该信息粒子加入到队列中。队列中的信息粒子均可以通过 EDF 传输策略进行发送,且可以证明此传输策略满足队列中所有信息粒子的时延要求。

2 粒子群的最小可达接入带宽分析

本小节给出了信息粒子及其属性的定义和信息粒子群及其属性的定义。定义了对信息粒子群中粒子的排序操作。定义了信息粒子群的最小可达接入带宽、最小可达接入带宽和平均接入带宽需求等。证明了 EDF 是可以达到最小可达接入带宽的接入传输策略。最后,给出了在给定信息粒子群的情况下,计算最小可达接入带宽的方法(注:更为完整的分析请参见文献[14])。

定义 1 信息粒子

具有特定基本属性(包括:承载量 $I_i \in \mathbf{R}^+$ 、初始时刻 $t_{b,i} \in \mathbf{R}$ 、截止时刻 $t_{e,i} \in \mathbf{R}^+$ 等)的待传输信息块被称为信息粒子,用 n_i 来标识。以下为其主要属性:有效生存区间为 $D_i \triangleq [t_{b,i}, t_{e,i}]$ (本文仅考虑已到达的粒子,即 $t_{b,i} = 0$), D_i 所覆盖的时间长度为此粒子的有效生存跨度 $T_{D_i} \in \mathbf{R}^+$,即 $T_{D_i} \triangleq t_{e,i} - t_{b,i}$ 。

信息粒子 i 在时刻 t 的瞬时传输速率记为 $r_i(t)$,若 n_i 在 D_i 时间区间内被发送,则称为有效发送。

定义 2 信息粒子群

包含 $N \geq 1$ 个有意义的信息粒子的非空集合,记为 Q 。其承载量为所包含的所有信息粒子承载量之和,即 $I_Q \triangleq \sum_{i \in Q} I_i$,其有效生存空间 D_Q 为所有信息粒子的有效生存区间的并区间,即 $D_Q \triangleq \bigcup_{i \in Q} D_i$,所涵盖的总时间长度为其有效生存跨度 T_{D_Q} 。

定义 3 对信息粒子群 Q 执行的升序排序操作 $O_{\text{inc}}(Q)$

给定信息粒子群 Q ,设其由 $N \geq 1$ 个信息粒子组成。如果按照各信息粒子截止时刻由小到大进行排序,并赋予排在第 1 位的信息粒子的序号为“1”(其截止时刻为 $t_{e,1}$),赋予排在第 2 位的信息粒子的序号为“2”(其截止时刻为 $t_{e,2} \geq t_{e,1}$),...,赋予排在第 N 位的信息粒子的序号为“ N ”(其截止时刻为 $t_{e,N} \geq t_{e,N-1}$),称这样的针对 Q 所包含的信息粒子进行的排序操作为升序排序操作,记为 $O_{\text{inc}}(Q)$ 。

定义 4 信息粒子群 Q 的可达接入带宽 $B_{\text{re},Q}$,单位:bit/s

如存在一种传输策略能使得在 Q 中的所有信息粒子都有效传输,且总的瞬时传输速率 $\sum_{i \in Q} r_i(t) \leq B_{\text{re},Q}$ ($B_{\text{re},Q} > 0, t \in D_Q$),则称接入带宽 $B_{\text{re},Q}$ 对 Q 来说是可达的。

定义 5 信息粒子群 Q 的最小可达接入带宽 $B_{\text{re},Q}^{\min} \in \mathbf{R}^+$,单位:bit/s

信息粒子群 Q 在其有效生存区间 D_Q 内的所有可达接入带宽中的最小值。

定义 6 信息粒子群的平均接入带宽需求 $B_{\text{av},Q} \in \mathbf{R}^+$,单位:bit/s

针对信息粒子群 Q ,定义 $B_{\text{av},Q} \triangleq I_Q/T_{D_Q}$ 为其平均接入带宽需求。平均接入带宽的实际物理含义为该粒子群中的所有粒子以恒定速率进行传输,且恰好在最大截止时间时完成对所有粒子传输的恒定速率,若这个恒定速率能使得该粒子群中的所有粒子有效发送,则其必为该粒子群的最小接入带宽。

定理 1 给定信息粒子群 Q ,设其最小可达接入带宽为 $B_{\text{re},Q}^{\min}$ 。对这个粒子群 Q 使用 EDF 策略可以在最小带宽下有效发送。

证明 首先,一个有效的发送方案需要满足的最基础条件为:在 $t_{e,i}$ 前完成对 n_i 的发送(参见定义

1 中粒子有效发送)。那么对 Q 实施递增排序操作 $O_{\text{inc}}(Q)$ 后,用 $Q_i(Q_i \subseteq Q)$ 表示信息粒子群 Q 中序号不大于 $i(1 \leq i \leq N)$ 的所有信息粒子组成的信息粒子子群。

条件 1 在 $t_{e,1}$ 时刻前的有效方案的最低要求为完成 Q_1 (即 n_1) 的发送,即:

$$B_{\text{re},Q}^{\min} \cdot T_{D_1} \geq I_1 \quad (1)$$

根据平均接入带宽的定义(定义 6),等价于

$$B_{\text{re},Q}^{\min} \geq B_{\text{av},Q_1} \quad (2)$$

条件 2 在 $t_{e,2}$ 前的最低要求为完成 Q_2 (即 n_1, n_2) 的发送,且满足条件 1,即

$$B_{\text{re},Q}^{\min} \geq B_{\text{av},Q_1} \text{ 且 } B_{\text{re},Q}^{\min} \geq B_{\text{av},Q_2} \quad (3)$$

等价于

$$B_{\text{re},Q}^{\min} \geq \max\{B_{\text{av},Q_1}, B_{\text{av},Q_2}\} \quad (4)$$

以此类推:

条件 N 在 $t_{e,N}$ 前的最低要求为完成 Q_N 的发送,且满足条件 $N-1$,即

$$\begin{aligned} B_{\text{re},Q}^{\min} &\geq \max\{B_{\text{av},Q_1}, B_{\text{av},Q_2}\} \geq \\ &\max\{B_{\text{av},Q_1}, B_{\text{av},Q_2}, B_{\text{av},Q_3}\} \geq \cdots \geq \\ &\max\{B_{\text{av},Q_1}, B_{\text{av},Q_2}, \cdots, B_{\text{av},Q_{N-1}}, B_{\text{av},Q_N}\} \end{aligned} \quad (5)$$

若 $\max\{B_{\text{av},Q_1}, B_{\text{av},Q_2}, \cdots, B_{\text{av},Q_{N-1}}, B_{\text{av},Q_N}\} = B_{\text{av},Q_k}$, 则 (5) 式取等号的条件为

$$B_{\text{re},Q}^{\min} \cdot T_{D_k} = \sum_{i=1}^k I_i \quad (6)$$

表示发送前 k 个信息粒子时完全利用了带宽为 $B_{\text{re},Q}^{\min}$ 的资源,且正好在 $t_{e,k}$ 时刻完成了传输。

所以

$$B_{\text{re},Q}^{\min} = \max_{i=1, \cdots, N}\{B_{\text{av},Q_i}\} \quad (7)$$

而使用 EDF 传输策略以瞬时速率 $B_{\text{re},Q}^{\min}$ 进行传输,则均可以对应每个条件,即使得每个粒子在其截止时间前完成发送,所以 EDF 是有效的。

证毕。

此定理给出了将整个粒子群有效发送的最低要求,且 EDF 方法正好满足所有的条件,即证明了 EDF 方法可以使用最小可达接入带宽有效发送整个粒子群。

定理 1 之推论 1:给定包含有限个(即 $N \geq 1$) 信息粒子的信息粒子群 Q ,对其执行递增排序操作 $O_{\text{inc}}(Q)$ 。令 B_{av,Q_i} 表示信息粒子群 Q 中序号不大于 $i(1 \leq i \leq N)$ 的所有信息粒子组成的信息粒子子群 $Q_i(Q_i \subseteq Q)$ 的平均接入带宽需求,则有

$$B_{\text{re},Q}^{\min} = \max_{i=1, \cdots, N}\{B_{\text{av},Q_i}\}。$$

证明 由定理 1 可知,在对 Q 实施递增排序操作 $O_{\text{inc}}(Q)$ 的前提下, $B_{\text{re},Q}^{\min} = \max_{i=1, \cdots, N}\{B_{\text{av},Q_i}\}$ 成立。

证毕。

此推论给出了对于已到达粒子所组成的粒子群,计算最小可达接入带宽的公式。

3 基于粒子化接入思想的细粒度延迟保障机制

本节首先基于粒子化接入思想,提出了一个基础性的细粒度延迟保障机制。进一步,为了改善网络重业务负荷情况下信息粒子间接入的公平性,提出了一种改进算法。

3.1 粒子化接入基础算法

如图 2 所示,不同种类的业务流信息粒子从上层到达后,都将与当前节点待发送队列中的信息粒子组成信息粒子群,然后进行准入判断,即计算最小可达接入带宽(参见定理 1 之引理 1),若最小可达接入带宽小于信道带宽,即认为该信息粒子能被有效发送,将其加入到队列中,反之则将其丢弃。队列中的信息粒子将等待 EDF 调度,即剩余时间越少的将越先被发送,且队列中的所有信息粒子都可以被有效发送(参见定理 1)。以下为粒子化接入基础算法。

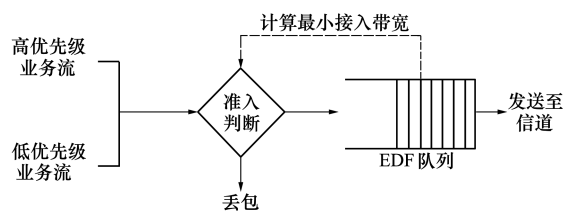


图 2 粒子化接入基础算法流程图

算法 1 粒子化接入基础算法

输入 到达的信息粒子

当前节点的待发送队列

输出 被发送的信息粒子

- 1 当一个信息粒子到达,计算其剩余时间
- 2 计算队列中粒子群以及到达粒子组成集合的最小可达接入带宽(参见定理 1 之引理 1)
- 3 IF (最小可达接入带宽 > 信道带宽)
- 4 将到达信息粒子丢弃
- 5 ELSE

```

6   将到达信息粒子加入节点 EDF 队列
7   END IF
8   WHILE (发送队列信息粒子群不为空)
9     选择发送队列中执行  $O_{\text{inc}}(Q)$  后排序第一的
      粒子(参见定义 3)
10  END WHILE

```

3.2 粒子化接入公平算法

根据粒子化接入最小可达带宽的计算公式可以得到,新加入的粒子的截止时间越小则计算得到的最小可达接入带宽越大,因此,在网络负载较重时,低时延的信息粒子会更难进入队列。为了平衡业务流之间的公平性,对算法 1 进行了改进并提出了粒子化接入公平算法,如图 3 所示,不同种类的业务流

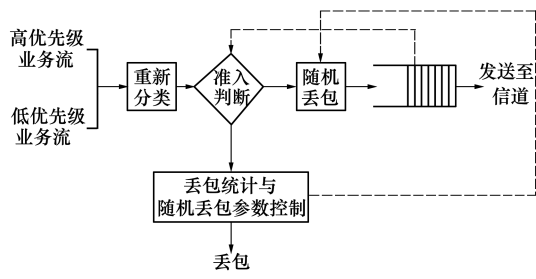


图 3 粒子化接入公平算法流程图

的信息粒子从上层到达时,首先根据前一段时间已经到达的信息粒子剩余时间的中值,将其分类为“紧急”或“一般”,然后再进行准入判断(同算法 1)。对通过准入判断的信息粒子增加一个随机丢包的环节。随机丢包阶段的目的是在高业务负载的状态下,保障 2 类信息粒子间的公平性。正常情况下“一般”类型的信息粒子时延容忍更高,因此其所需的最小可达接入带宽更小,更容易进入队列并获取到接入资源。为公平性起见,需要控制“一般”类型信息粒子的接入成功率。具体做法是,在“一般”类型信息粒子通过准入判断后,依旧会存在一定概率被随机丢弃,这样可以使得队列腾出资源给“紧急”类型的信息粒子。首先定义平衡阈值和固定丢包率,平衡阈值为“紧急”类型的业务占比目标。固定丢包率为当实际发送“紧急”类型业务占比偏移于平衡阈值时,调整“一般”类型丢包率的单次调整量,具体的调整过程如下:统计一定时间内,丢弃的信息粒子类型,若“一般”类型的丢弃占比超过平衡阈值则降低固定丢包概率,若“紧急”类型的丢弃占比超过平衡阈值则增加固定丢包概率。这种反馈机制和重新分类的方法可以改善不同类型信息粒子接

入的公平性。最后在进入队列中的信息粒子将等待 EDF 调度和发送。以下为改进算法的具体描述:

算法 2 粒子化接入公平算法

```

输入  到达的信息粒子
        当前节点的待发送队列
输出  被发送的信息粒子
1  当一个信息粒子到达,计算其剩余时间
2  IF (剩余时间<统计的剩余时间中值)
3    将该粒子分类为“紧急”
4  ELSE
5    将该粒子分类为“一般”
6  END IF
7  更新统计的时间中值
8  计算队列中粒子群以及到达粒子组成的集合的
   最小可达接入带宽(参见定理 1 之引理 1)
9  IF(最小可达接入带宽>信道带宽)
10   将到达信息粒子丢弃,更新丢包统计
11 ELSE
12   进入随机丢包阶段
13 END IF
14 IF (到达粒子为“一般”且 random[0,1)>“一
   般”固定丢包率
15   将到达信息粒子丢弃,更新丢包统计
16 ELSE
17   将到达信息粒子加入发送队列中,进行传输
18 END IF
19 更新丢包统计:计算前 100 个丢包中“一般”分
   类包的占比
20 IF (“一般”分类占比>平衡阈值)
21   固定丢包率=固定丢包率+丢包率步长
22 ELSE IF (“紧急”分类占比>平衡阈值)
23   固定丢包率=固定丢包率-丢包率步长
24 ELSE
25   固定丢包率不变
26 END IF
27 WHILE (发送队列信息粒子群不为空)
28   选择发送队列中执行  $O_{\text{inc}}(Q)$  后排序第一的
      粒子(参见定义 3)
29 END WHILE

```

4 仿真验证

本文针对无线网络中的下行场景进行仿真,如图 4 所示,中心节点 (AP) 需要同时传输 n 条第一类业务流和 m 条第二类业务流。每条业务流的到达强度相同且满足泊松分布,在仿真实验中,业务流的到达强度用每毫秒发送分组个数来描述。为了模拟实际情况中不同类型业务流的特征,每类业务流中包含 2 种不同时延容忍分布的信息粒子。第一类业务流中,低时延信息粒子占比较高,但也存在少量高时延信息粒子;第二类业务流中则恰好相反,存在大量的高时延信息粒子和少量低时延信息粒子。为简单起见,业务流中各类信息粒子的分布都采用均匀分布。

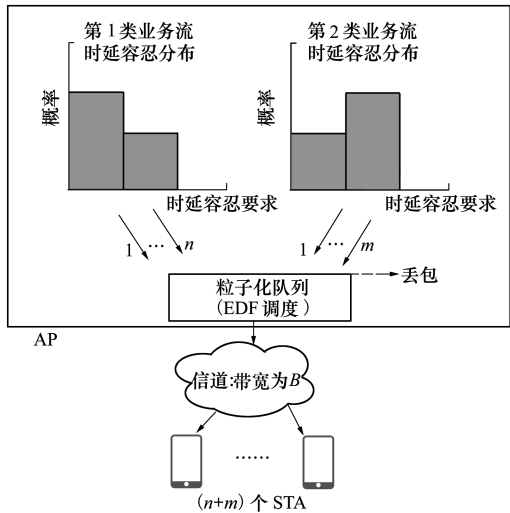


图 4 粒子化接入方案仿真示意图

4.1 粒子化接入方案的仿真实验描述

如图 4 所示,本文的实验方案中,2 类业务流都将执行粒子化接入公平算法(参见算法 2),即首先对信息粒子进行准入控制,其次对准许进入队列的信息粒子按照 EDF 策略依次进行传输。

4.2 对比方案的仿真实验描述

本文的对比方案是当前常用“基于类”的接入机制的简化实现版,如图 5 所示。与 4.1 中实验方案相同的 2 类业务流到达节点后,各自进入相应的 FIFO 队列中进行排队等待。在信道接入环节,首先设置队列的发送权重,调度器将按照发送权重占比来随机调度发送不同优先级队列中的信息粒子。

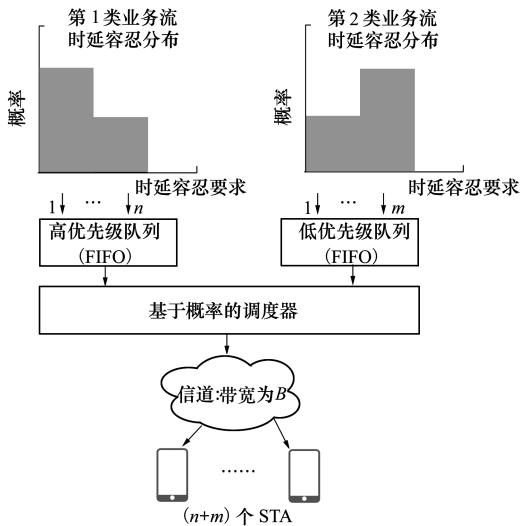


图 5 仿真对比方案示意图

4.3 仿真环境及参数设置

仿真实验中本文利用 C++搭建仿真平台,其中应用层按照设定参数和分布随机产生 2 类业务流,然后数据链路层分别使用粒子化接入方案和对比方案对其进行处理,发送和丢包数据在信道层被统计。

表 1 仿真参数表

仿真参数		数值
仿真时段/ms		0~3 000
信道带宽/MHz		40
平衡阈值/%		50
固定丢包率步长		0.02
第一类业务流	低时延粒子占比	0.9
	高时延粒子占比	0.1
	业务流数量 n	10
第二类业务流	低时延粒子占比	0.1
	高时延粒子占比	0.9
	业务流数量 m	10
低时延信息粒子	承载信息量/bit	12 000±200
	时延容忍/ms	50±10
高时延信息粒子	承载信息量/bit	12 000±200
	时延容忍/ms	500±100
对照队列	高优先级队列	8
发送权重	低优先级队列	4

本文在保持其他参数不变的情况下,逐渐增加每条泊松业务流的到达强度,对比 2 种方案的有效吞吐量(指单位时间在其截止时刻内完成有效发送的信息粒子/分组的总比特数)、平均时延(指所有信息粒子/分组从其产生到完成发送的平均时间)、

以及丢包率(指未有效发送的信息粒子/分组占总产生粒子/分组数的比值)。仿真的具体参数如表 1 所示,其中,平衡阈值和固定丢包率只在算法 2 中被使用,详见 3.2 节。对照队列发送权重只在对照组中被使用,详见 4.2 节。其他参数均采用 WiFi 环境的经典设置,为了仿真方便,本文仿真中采用无差错理想信道。

4.4 仿真结果及分析

如图 6 所示,在业务流负载较轻的情况下,3 种接入方法的性能非常接近,但随着业务到达强度的增加,对比方案由于没有进入队列的准入控制,队列中的分组出现堆积。而且已经失效的分组依然需要占用信道资源,这更加使得后续的分组无法得到有效发送,从而使对比方案的有效吞吐量急剧下降。而本文提出了粒子化接入方案由于其入队之前便可计算出到达的分组能否被有效发送,无法被有效发送的分组将被直接丢弃,这样可以节约有限的信道资源。所以,如图 6 所示,2 种粒子化接入方案在业务负载较高的情况下,都依然能保持较高的有效吞吐量,几乎都能完全充分地利用整个信道(仿真信道带宽 40 MHz)。另外,由图可见,在业务负载逐渐增加时,粒子化接入基础算法 2 类业务流的有效吞

吐量差距逐渐增大,但粒子化接入公平算法的 2 类业务的有效吞吐量依旧能保持接近,可见相比于粒子化接入基础算法,粒子化公平算法在公平性上有较好改善。

从图 7 来看,对比方案大部分信息粒子发送完成后都已经失效,即超过了其截止时间(本文仿真中最大的截止时间设为 600 ms),这是其有效吞吐量急剧下降的原因。而粒子化接入方案由于具有准入控制,保证了进入队列的信息粒子的传输都是有效的,且较好地避免了信道传输资源的浪费,因此平均时延性能远优于对比方案。而在粒子化接入算法中,由图可见在粒子化基础算法中低时延占比较高的业务(第一类)不容易被发送,因此其平均时延相比公平算法也相对较高。

从图 8 中可以看出,对比方案在业务流到达强度逐渐增加到饱和的情况下,丢包率会突增至 80% 以上,此时接入网络几乎瘫痪。而粒子化接入的丢包率则呈现较为平缓的上升趋势,在负载较重的状态下,依然能保持较低的丢包率。并且图中可以看出,粒子化基础方案中 2 类业务流丢包率差距明显,存在公平性问题,而粒子化公平方案有效解决了这个问题,2 类业务流的丢包率一直比较接近。

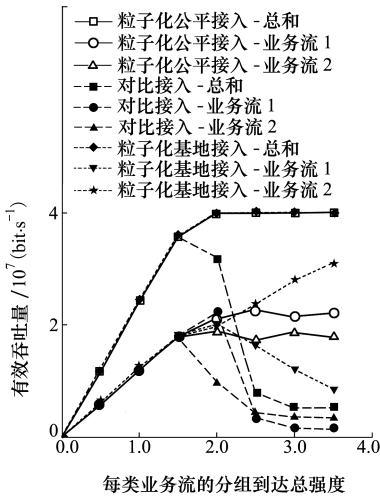


图 6 粒子化接入及对比方案各业务流有效吞吐量曲线

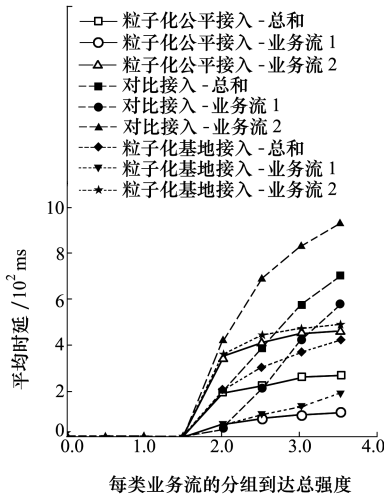


图 7 粒子化接入及其对比方案各业务流平均时延曲线

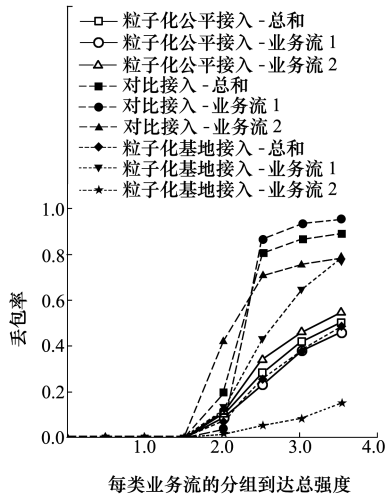


图 8 粒子化接入及其对比方案各业务流丢包率曲线

综上,粒子化接入机制相比于传统的“基于类”方法,在业务流负载较轻的情况下性能相似,但在业务流负载很重的情况下,传统的“基于类”方法会由于队列中的分组堆积导致大部分分组不能得到及时

发送,使得网络瘫痪。而粒子化接入策略由于在分组到达时就可以通过计算最小可达接入带宽来判断该分组是否能被有效发送,不会在队列中堆积,所以在高业务负载的状态下性能远超传统方案,且得到

发送的分组都能得到相应的时延保障。进一步,从吞吐量和丢包率性能对比来看,粒子化接入公平方案有效解决了粒子化接入基础方案中不同业务流公平性较差的问题。

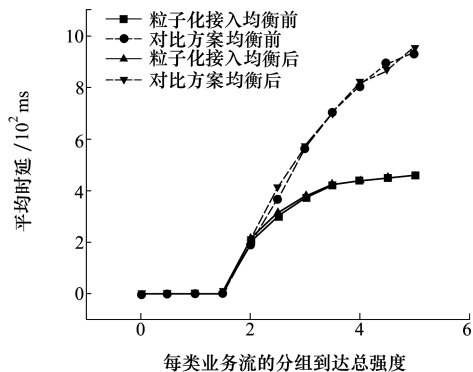


图 9 粒子时延类型占比不同情况下粒子化接入和对比方案平均时延对比

为了模拟业务流中 2 类分组占比更均衡的场景,实验将第一类业务流的高/低时延分组占比从 1:9 修改为 3:7;第二类业务流的高/低时延占比从 9:1 修改为 7:3。仿真结果为图 9 所示,本文提出的粒子化接入算法依然明显优于对比方案。对比方案由于不关心业务流的内部情况,因此性能没有改变。而粒子化接入方案由于两类业务流总和的高/低时延粒子比例修改前后保持不变,且方案不关注信息粒子来源,因此性能没有变化。

参考文献:

- [1] NOSHEEN Summera, KHAN Jamil Yusuf. Quality of service-and fairness-aware resource allocation techniques for IEEE 802.11ac WLAN[J]. IEEE Access, 2021, 9: 25579-25593
- [2] PISCHELLA Mylene, RUYET Didier Le. NOMA-relevant clustering and resource allocation for proportional fair uplink communications[J]. IEEE Wireless Communications Letters, 2019, 8(3): 873-876
- [3] LU Cong, WU Bin, YE Tianchun. A novel QoS-aware a-MPDU aggregation scheduler for unsaturated IEEE 802.11n/ac WLANs[J]. Electronics, 2020, 9(8): 1203
- [4] LYU Yunxin, DIAS Maluge Pubuduni Imali, RUAN Lihua, et al. Request-based polling access: investigation of novel wireless LAN MAC scheme for low-latency e-health applications[J]. IEEE Communications Letters, 2019, 23(5): 896-899
- [5] COSTA Robson, LAU Jim, PORTUGAL Paulo, et al. Handling real-time communication in infrastructured IEEE 802.11 wireless networks: the RT-WiFi approach[J]. Journal of Communications and Networks, 2019, 21(3): 319-334
- [6] LI Bo, LI Wei, VALOIS Fabrice, et al. Performance analysis of an efficient MAC protocol with multiple-step distributed in-band channel reservation[J]. IEEE Trans on Vehicular Technology, 2009, 59(1): 368-382
- [7] LI Bo, MOHAMMED G A A, YANG Mao, et al. The design methodology for mac strategies and protocols supporting ultra-low delay services in next generation IEEE 802.11 WLAN[C]//International Conference on Internet of Things as a Service, Cham, 2020: 80-97
- [8] FARAG Hossam, SISINNI Emiliano, GIDLUND Mikael, et al. Priority-aware wireless fieldbus protocol for mixed-criticality in-

5 结 论

本文针对无线网络中多业务的时延保障问题,提出了粒子化接入的思想,并基于此提出了一种无线网络中的细粒度延迟保障机制。相比与传统的时延保障机制,本文提出的细粒度延迟保障机制所考虑的接入机制操控对象更加基础,即分组级别,并且所考虑的操控对象并不关注其来源和优先级,只考虑信息粒子的截止时间和承载信息量,因此具有更细的服务质量保障粒度和更好的可拓展性。进一步发现,在网络负载较重时,若信息粒子的截止时间很短,则在粒子化接入算法中,这个粒子将更不容易通过网络,这样的机制鼓励用户如实标记信息粒子的截止时间。在该机制下,如果网络比较拥塞,用户还可以选择在容忍范围内尽可能增大信息粒子的截止时间,来换取更高的网络通过性。可见,从根本上来讲,基于粒子化接入思想的细粒度延迟保障机制鼓励接入用户以更为诚实而高效的方式使用网络接入资源。另外,由于粒子化接入思想所具有的一般性,其还可被应用于改善包括 Internet 网在内的其他类型网络的服务质量保障性能。

下一步工作中,将进一步研究如何将本文所提出的接入机制拓展至更为一般的应用场景,如上行应用场景和复杂传输信道场景等。

- dustrial wireless sensor networks[J]. IEEE Sensors Journal, 2018, 19(7): 2767-2780
- [9] YI Changyan, CAI Jun. A truthful mechanism for scheduling delay-constrained wireless transmissions in IoT-based healthcare networks[J]. IEEE Trans on Wireless Communications, 2018, 18(2): 912-925
- [10] KARADAG Goksu, GUL Recep, SADI Yalcin, et al. QoS-constrained semi-persistent scheduling of machine-type communications in cellular networks[J]. IEEE Trans on Wireless Communications, 2019, 18(5): 2737-2750
- [11] IQBAL Adeel, HUSSAIN Riaz, SHAKEEL Atif, et al. Enhanced spectrum access for QoS provisioning in multi-class cognitive D2D communication system[J]. IEEE Access, 2021, 9: 33608-33624
- [12] SHAH Sayed Qaiser Ali, KHAH Farrukh Zeeshan, BAIG Adeel, et al. A QoS model for real-time application in wireless network using software defined network[J]. Wireless Personal Communications, 2020, 112(2): 1025-1044
- [13] TANENBAUM A S, WETHERALL D J. Computer networks[M]. 5th ed. Upper Saddle River, New Jersey: Prentice Hall, 2011
- [14] LI Bo, SUN Ke, YAN Zhongjiang, et al. Idea and theory of particle access[EB/OL]. (2012-03-29)[2022-05-20]. <https://arxiv.53yu.com/abs/2203.1519/>

Study on fine-grained mechanism of delay guarantee based on idea of particle access in wireless networks

SUN Ke, LI Bo, YAN Zhongjiang, YANG Mao

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Aiming at the problems existing in the current mechanisms of delay guarantee in wireless networks (i.e. poor scalability, coarse granularity for provided service levels, and improving delay performance at the expense of sacrificing some resource utilization), this paper puts forward both the idea of particle access and the corresponding access mechanism. In this paper, a traffic flow is modeled as a group of information particles that carry a certain amount of information and are valid for certain periods of time. Firstly, the definitions of information particles and the group of information particles are given. It is proved that the minimum reachable access bandwidth of an information particle group can be achieved by using the EDF (earliest deadline first) transmission strategy. Moreover, a fine-grained mechanism of delay guarantee based on the idea of particle access is proposed for the dynamic access environment in wireless networks. Extensive simulations are carried out for the application scenario of downlink transmissions, and it is shown that, in the case of heavy traffic loads, comparing with the rapid growth of the average packet delays and the packet loss rates in a legacy access mechanism, the proposed fine-grained access strategy based on the idea of particle access can also achieve the better performance on the average packet delays and the packet loss rates, and hence higher effective throughput is obtained. The research in this paper pave a new way for further improving quality of service (QoS) mechanisms of wireless networks.

Keywords: wireless networks; access mechanism; particle access; delay guarantee

引用格式: 孙珂, 李波, 闫中江, 等. 无线网络中一种基于粒子化接入思想的细粒度延迟保障机制研究[J]. 西北工业大学学报, 2023, 41(1): 81-89

SUN Ke, LI Bo, YAN Zhongjiang, et al. Study on fine-grained mechanism of delay guarantee based on idea of particle access in wireless networks[J]. Journal of Northwestern Polytechnical University, 2023, 41(1): 81-89 (in Chinese)