**蜂窝网M2M通信中基于随机网络演算的接纳控制建模与性能分析**

**1 引言**

物联网的迅速发展给智慧网络不仅带来了新的发展机遇，也带来了新的挑战。M2M通信作为物联网的重要组成部分，旨在减少人为干预实现机器与机器之间自主通信，在广义上又包括了人与机器及机器与人的通信形式。M2M通信可以选择任何网络通信技术作为其传输数据的通道，如互联网、有线电信网、蜂窝网等，而由于蜂窝网的无线覆盖范围广，且使用便利，在实现M2M通信时具有较大优势。3GPP将此类通过蜂窝网作为M2M数据传输的方式定义为蜂窝网M2M通信[1]。

蜂窝网中的M2M设备呈爆炸式增长，因此建立一个良好的物联网生态环境，不仅要考虑如何处理急剧增加的设备的大规模连接，还要保障每个M2M设备数据传输的QoS需求，许多文献针对M2M通信存在的问题进行了研究。文献[2]总结了M2M网络近年来的发展并提出了面临的挑战，这些挑战包括了大规模设备的维护和远程管理问题，[3]给出了这些问题的解决方案。[4]和[5]研究了如何在LTE网络中进行资源管理优化以支持M2M通信的问题。容易发现，现有的M2M解决方案大致可以分为两类：（1）无线资源管理优化（2）设备间的协同工作。其中，无线资源的管理优化在保障QoS方面是至关重要的，针对此问题，[6]提出了可靠的资源池方案，[7]引入接入限制参数来改变资源块从而优化系统性能，[8]提出了批量处理数据模型以减少M2M网络的更新频率。另外，为了支持大量的机器设备接入，可以通过对大量设备进行分组，利用分组中各设备的相互协作实现负载均衡和资源共享，比如[9]分析了分簇M2M网络，研究了在LTE-A环境下随机接入资源的空间复用方法，此方法不但支持大量M2M设备的接入，并为H2H通信预留了更多的随机接入资源。

综上所述，大多数M2M相关研究关注于资源管理，基站负载均衡以及对M2M设备分组等方面，而少有文章通过解决M2M请求的接纳控制问题来优化整个通信网络。针对蜂窝网下大量M2M设备并发接入的场景，建模其接纳控制问题具有一定难度，需要使用更加复杂的能透彻分析该问题的工具。网络演算作为迅速发展起来的性能分析理论，不用考虑网络是否达到稳态以及流量的复杂性，拥有一套系统的方法分析网络性能，成为解决通信系统建模和性能分析的强大手段。相比于传统的排队论，网络演算使用包络的方法描述系统接收的数据和系统的服务能力，无需关心数据流的具体特性以及系统的具体实现方式，所以能够为各种网络场景提供通用的建模与分析方法。

现有文献[10]使用了该理论处理M2M请求的接纳控制问题，但是使用了确定性网络演算，没有考虑现实网络中大量流量带来的统计特性，浪费了网络资源，因而本文将使用随机网络演算进行接纳控制算法的分析，同时将M2M请求和蜂窝请求一同考虑进来，更加符合现实的物联网通信。

**2 M2M接纳控制模型**

**2.1 系统模型**

M2M通信越来越多样化和复杂，需要一个开放、完善的架构来支持它，而现有的IoT网络还无法提供一个统一化、标准化的架构以包容M2M通信和更多可能出现的通信服务，因此这里提出一个更便于管理和控制整个通信网络的IoT架构，如图1所示。



图1基于SDN的IoT通信架构图

上图的网络架构中将传统的网络架构分解成了由控制层、接入层和设备层组成的三层架构，每一层的作用说明如下。

设备层：该层为整个系统架构的最底层，主要组成元素是各种无线终端，包括M2M设备。

接入层：处于系统架构的中间层，一方面通过蜂窝或者wifi等技术接收最底层终端接入的连接信息，另一方将接入请求和相关数据传输给上层，作为上层控制决策的信息依据。

控制层：即SND控制器。主要负责特定策略算法定制以及下发指令等，本课题主要考虑的是大量M2M设备的接纳控制算法的制定以及实施。SDN下的控制器通过南向接口与各基站或AP点相连，通过站点回传的相关信息进行统一的实时监控，并执行接纳策略，下发执行指令对站点进行控制。

该架构中将网络控制和数据的传输分离开来，传承了SDN的设计思想，提高了管理和控制网络的便利性和灵活性。同时，控制层与接入层之间的交互行为与SDN相关协议规定的行为保持一致，本文接下来的研究都是在假设实施了该行为标准后进行的。

**2.2 接纳控制模型**

本文将重点分析不同类M2M设备及蜂窝用户建立通信连接而发送请求时接纳控制过程的性能评估，而不是数据转发过程的性能评估，即我们关注于SDN体系中的控制层，而非数据层。在图1基于SDN的IoT网络架构下，为了描述各类通信请求接入到基站后的接纳控制问题，本文提出一种基于多优先级的接纳控制模型，如图2所示。其中，每个基站将接收到的所有请求分类为四种优先级队列，分别为用于存放蜂窝用户请求、M2M实时请求、M2M非实时请求以及M2M时延容忍请求。当一个请求进入基站点时，首先根据请求类型将其归类为对应的四种请求之一，然后被路由到对应等级的队列中等待服务。



图2 蜂窝网M2M通信中基于多优先级队列的接纳控制模型

服务器现有优先级为的请求正在被处理，则服务器接下来的服务顺序满足：a.优先级低于的的M2M设备请求服务时，服务器将服务完当前的请求流，再为低优先级请求提供服务；b.优先级高于的M2M设备请求服务时，服务器将在服务完请求的当前数据后按优先级依次为其他M2M请求提供服务；c.若蜂窝用户请求到达时，系统将停止当前的一切服务转而为蜂窝用户请求提供服务。可见，该模型中不仅包括了M2M请求的非抢占式调度，还包括了蜂窝用户请求对M2M请求的抢占式调度，有助于更加全面地分析蜂窝网M2M通信请求的接纳控制问题。

此外，该模型完全可以处理其他的场景，比如，在需要传输紧急信息的场景下，会出现M2M请求的优先级高于蜂窝用户请求的情况，并且针对不同的应用场景，区分M2M请求实时、非实时、时延容忍的标准是不一样的，然而该问题不在本文研究范围内，但确实有待进一步研究。接下来，我们将分析出该模型中每种请求流的性能，深入了解接纳控制的性能结果。

**3 性能分析**

本章节将首先介绍本文计算会涉及到的随机网络演算相关知识，接着将该理论应用到该模型中，分析该模型下的常见性能参数。具体地，我们将研究各类请求流对应的随机服务曲线、时延以及积压。

**3.1 随机网络演算基本知识**

网络演算分为两个分支：确定网络演算和随机网络演算。确定网络演算分析的是最坏情况下的网络性能，而这种最坏情况发生的概率比较小，因此分析结果不具有实用性。本文使用的随机网络演算引入了概率运算，可以描述和分析网络数据流量的统计复用特性，充分利用网络资源，弥补了确定网络演算的不足，下面将列出随机网络演算的相关运算知识。

在网络演算中，数据流的到达过程、服务过程和离开过程分别用、和描述，表示时间内到达数据流、获得服务的数据流以及离开数据流的累积量，三者都是关于时间的随机过程。

**定义1（随机到达曲线）** 若对于数据流存在和，使得对于所有的和满足

 (1)

则称该流具有v.b.c随机到达曲线，为到达数据流的流量上界函数，为到达曲线的概率上界函数，记为。

**定义2（随机服务曲线）**对于数据流经过系统S获得服务后离开过程为，若存在和，使得对于所有的和满足

 (2)

则称该服务系统S为数据流提供随机服务曲线，为系统S所提供的服务下界函数，为服务曲线的概率上界函数，记为，运算符代表了最小加卷积，即。

在已知到达数据流满足的随机到达曲线以及系统提供服务的随机服务曲线之后，可以得到如下性能定理。

**定理3（性能边界）**满足v.b.c随机到达曲线的输入流进入服务系统S获得服务，则对于所有的和，可得到时延边界和流量积压边界

**** (3)

**** (4)

其中代表和间的最大水平距离，表示和间的最大垂直距离。

**3.2 随机到达曲线的分析**

为了所有接入请求能够被平滑处理，我们假设采用令牌桶算法调节所有接入请求。令牌桶是一种典型的用于限制数据速率的整形算法，其基本工作机制如下：一个固定大小为的桶，每秒中有个令牌加入到该桶中，每当一个长度为的数据流发送到桶中，则将删除桶中的个令牌，并将流发转发到网络中，而当桶中的令牌不足时，数据流将被放入队列中等待桶中的令牌个数满足要求时，再删除个令牌并转发当前数据。

由此可见，任何输入流经过令牌桶整形后，得到的数据流的随机到达曲线表示为，且它受限于。其中是数据流的平均传输速率，****是数据流的最大突发量，是任意不小于0的自由参数。

**3.3 随机服务曲线的分析**

已知各类请求的随机到达曲线后，为了分析它们在网络中的传输性能，就要重点分析各类请求对应的随机服务曲线。

我们假设用表示蜂窝用户的请求流，表示优先级为的M2M请求流，其中如果，则称的优先级高于。而基站的总服务能力使用速率-延迟函数[11]表示，即基站的服务能力满足，其中。另外，用表示请求流的最大长度，表示优先级低于i流的M2M请求流中的最大流长度。为了不失一般性，这里举一个典型的例子来说明服务系统为不同优先级流的服务顺序，以便分析随机服务曲线，如图3所示。



图3 不同优先级流获得服务的顺序说明图

其中，、、表示M2M请求优先级从高到低的请求流，分别对应实时、非实时、时延容忍三类请求，表示最高优先级的蜂窝网请求。、、、分别是、、、四种流到达系统并开始产生积压的时刻。在该例中，最先到达基站，所以基站将为其首先提供服务。当系统正在服务时，和也依次到达基站，则基站将会在服务完后依次服务和。紧接着，当在系统服务的过程中到达时，系统将立即终止对的服务，转向服务，等处理完后再继续服务。

假设所有队列在零时刻长度为空，数据包最后一比特到达系统才能获得服务，并且同一数据流中的包按照FIFO规则依次被处理。下面我们首先分析的性能，时间内的系统总输出流为，则

 (5)

由于在特定的时间点，输出流量等于输入流量，但对于变化的时间，，所以

 (6)

同理，

 (7)

 (8)

通过以上式子，得到

 (9)

其中

 ， (10)



将式（8）变形为

 (11)

又根据定义2，上式可写成

 (12)

由此可以看出如果是广义递增函数，则系统S为提供的服务满足。

接下来分析,在时间内的系统总输出流为,则

 (13)

类似分析的方法，得到

 (14)

即系统S为提供的服务满足，其中。

同理得到，系统为提供服务，其中。

通过以上分析，发现各类M2M请求的随机到达曲线满足形式，又根据3.2节分析可知，各类M2M请求的随机到达曲线可以表示为，所以

 (14)

 (15)

 (16)

由此可见，整个服务系统提供给M2M请求的服务能力下界不仅与蜂窝用户请求以及优先级高于当前流的M2M请求流有关，还与低于当前请求流的M2M请求流的最大请求长度有关。

因此，我们可以就可以计算出时间内的累积输出:

 (17)

从而

 (18)

其中，如果如果是广义递增函数，则系统提供给M2M请求流的服务满足，而且

 (19)

此外，由于蜂窝用户请求对于所有M2M请求是抢占式的，即系统会将所有服务提供给蜂窝请求，无论当前的请求是否处理完毕。所以,且系统为提供服务，其中。

**3.4 性能分析结果**

现在根据定理1进行性能分析，我们首先利用式（3）计算时延边界。其中

****（20）

所以M2M请求的最大时延满足

**** （21）

同样，对于蜂窝用户请求，其

 （22）

因此蜂窝用户请求的最大时延满足

**** （23）

根据式（21）和式（23），可以发现M2M请求的时延受多个因素的影响，其中以高于自身优先级的请求流特性为主，而蜂窝用户请求的时延只和、有关。接下来，再关注一下它们的积压长度和哪些因素有关。

利用式（4），其中

 (24)

因此M2M请求经过系统的积压长度满足

**** (25)

类似地,

(26)

因此蜂窝用户请求经过系统的积压长度满足

**** (27)

以上对系统总积压长度的分析显示出，M2M请求流的积压不仅与自身的随机到达曲线的特性有关，还与比自身优先级高的请求流有关，而蜂窝用户请求的积压长度只与它自身有关。

**4 数值分析**

这一小节将会评估上一节计算得到的性能边界，通过大量数值分析探讨影响性能边界变化的因素，以及具体的影响情况。以下分析中，我们设置默认相关参数值如下：,,各类请求流的最大流长度为。

图4展示了系统提供给各类请求流的最小服务速率与各流到达速率的关系。图4中(a)~(d)分别对各类请求流的到达速率做了如下假设：(a)，；(b)，;(c),;(d),。我们可以发现低优先级流的最小服务速率总是受比其高优先级的流到达速率的不断增大而变小。例如，图4(c)中到达速率不断变大，影响的只是和这两种比优先级低的流的最小服务速率，而它自身和对应的最小服务速率确实不变的，这样的变化趋势也与式(14)~(16)的数学分析结果一致。

****

图4 最下服务速率随到达速率的变化

****

图5 概率边界随时延边界的变化

****

图6 时延边界与各流到达速率的关系

图5给出了概率边界与各类流时延边界的关系。此处的分析设置了所有流的到达塑料厂相等，即。很明显，概率边界越大，对应的时延边界越小，并且当固定概率边界值时，各类请求流的时延边界按优先级从高到低的顺序依次变大，符合式（3）时延边界满足的定理。为了分析不同流的时延边界和各类流的到达速率之间的关系，我们固定自由参数为某一合理值，，给出了如图6中的结果。观察图6(a)~(d),发现无论各类流的到达速率变化如何，蜂窝用户请求流的时延边界总是比M2M请求流的时延边界更小，且稳定为某一个值，这说明了蜂窝用户的请求能够得到更高的QoS保障。此外，在M2M请求流中，高优先级流的时延低于低优先级流，更重要地，随着到达速率的不断变大，各流之间的时延差距越来越大。以图6(d)为例，当不断增大的到达速率时，、、的时延边界之间的差距也在变大，且的增长趋势最快，即随着到达速率的不断增大，优先级越低的流受影响的程度越大。

图7描述了概率边界与各类流积压长度边界的关系。和图5类似，概率边界越大，对应的积压长度边界越小，并且当固定概率边界值时，各类请求流的积压长度边界按优先级从高到低的顺序依次变大，符合式（4）积压边界满足的定理。接着，图8分析了各类流的到达速率对最大积压长度的影响情况。从图8(a)~(c)看出，到达速率的变化不仅影响自身的积压长度，同时影响了比其优先级更低的流的积压长度。图8(d)中只有的积压长度不变，其他M2M请求流的积压长度都在变。



图7 概率边界随积压长度的变化



图8 积压长度边界与各流到达速率的关系

**5 本章小结**

本章节针对M2M中海量并发接入请求场景，选用了更合适的随机网络演算方法去分析M2M的接纳控制算法，弥补了现有文献选用确定网络演算解决M2M接纳问题的缺点，同时将蜂窝网中传统的请求和M2M请求同时考虑进来，使得分析结果更接近实际网络。其中，讨论的接纳控制模型符合多优先级队列模型，将所有请求分为蜂窝用户请求、M2M实时请求、M2M非实时请求以及M2M时延容忍请求这4类，最后分析了它们在该接纳控制模型中各自的时延和节点积压长度的边界值。

**校对报告**

当前使用的样式是 [Numbered(Multilingual)]

当前文档包含的题录共9条

有2条题录存在必填字段内容缺失的问题

参考文献 [5] ：字段(会议地点)内容缺失;

字段(出版社)内容缺失;

参考文献 [9] ：字段(会议地点)内容缺失;

字段(出版社)内容缺失;

**References:**

[1]. Taleb, T. and A. Kunz, Machine type communications in 3GPP networks: potential, challenges, and solutions. Communications Magazine IEEE, 2012. 50(3): p. 178 - 184.

[2]. Chen, M., et al., A Survey of Recent Developments in Home M2M Networks. IEEE Communications Surveys & Tutorials, 2014. 16(1): p. 98-114.

[3]. Taleb, T. and A. Kunz, Machine type communications in 3GPP networks: potential, challenges, and solutions. Communications Magazine IEEE, 2012. 50(3): p. 178-184.

[4]. Lien, S.Y. and K.C. Chen, Massive Access Management for QoS Guarantees in 3GPP Machine-to-Machine Communications. IEEE Communications Letters, 2011. 15(3): p. 311-313.

[5]. Matamoros, J. and C. Antón-Haro. Data aggregation schemes for Machine-to-Machine gateways: Interplay with MAC protocols. in Future Network & Mobile Summit. 2012.

[6]. Madueño, G.C.,Stefanović and P. Popovski, Reliable Reporting for Massive M2M Communications With Periodic Resource Pooling. IEEE Wireless Communications Letters, 2014. 3(4): p. 429-432.

[7]. Wiriaatmadja, D.T. and K.W. Choi, Hybrid Random Access and Data Transmission Protocol for Machine-to-Machine Communications in Cellular Networks. IEEE Transactions on Wireless Communications, 2015. 14(1): p. 33-46.

[8]. Sou, S.I. and S.M. Wang, Performance Improvements of Batch Data Model for Machine-to-Machine Communications. Communications Letters IEEE, 2014. 18(10): p. 1775-1778.

[9]. Wang, S.H., et al. Random access design for clustered wireless machine to machine networks. in First International Black Sea Conference on Communications and NETWORKING. 2013.

[10]. Huang, J., et al., Optimizing M2M Communications and Quality of Services in the IoT for Sustainable Smart Cities. IEEE Transactions on Sustainable Computing, 2017. PP(99): p. 1-1.

[11]. Stiliadis, D. and A. Varma, Latency-rate servers: a general model for analysis of traffic scheduling algorithms. IEEE/ACM Transactions on Networking, 1998. 6(5): p. 611-624.