**一 引言**

云计算采用集中式控制，将各种虚拟化的资源统一的存储在云数据中心，云数据中心通过网络，为网络中的用户提供各式各样的服务，以满足网络中用户的请求。近几年，智能手机、可穿戴设备和车载终端等智能移动终端爆发式的增长，无线网络传感和传输技术的提升，无线网络覆盖区域进一步扩大，现有的云计算框架可满足大多数固定的网络用户（PC端用户），但无法满足移动互联网时代的智能移动终端的多样化服务需求。为了适应新需求，思科首先提出了雾计算的概念。雾计算的出现，综合考虑移动智能终端的各种需求特性，按照提供本地化服务的思想，根据智能终端的行为特征和消费需求，通过预测提前将部分数据存储到靠近用户的雾计算边缘存储节点上，实时的处理终端用户的请求，有效的弥补云计算框架的不足之处。然而雾计算虽有相应的计算能力和存储能力，但是雾计算的边缘存储节点缓存空间大小相比云计算很有限，为了满足移动终端用户的需求，为此有必要选择性的将部分终端用户经常请求的数据存储到本地的雾计算边缘存储节点，并且当缓存空间不足的时候能置换出边缘节点中的某些数据对象，使得雾计算节点存储的可以保持最佳的存储状态。

雾计算作为一种新型的侧重于为移动智能终端用户提供服务的新型网络计算模式，近几年开始在国内逐渐的兴起。计春雷，杨志和，谢致邦等在文献[13]介绍了雾计算产生的时代背景、基本概念、雾计算的分层体系和分类，并用集合论和场域论等数学知识形象化的描述雾计算各个要素彼此之间的关系。程冬梅和李志等在文献[14]首先提出一个通用的雾计算框架。然后研究介绍雾计算的关键技术和主要特点，特别是针对医院的场景，设计出能提供患者就医相关信息的浏览服务，比如排队时长的查寻、医药价格的查询等，通过实际的实验操作验证了雾计算系统的实用性和可靠性。Jingtao，Fuhong，ZHOU等在文献[15]研究了雾环境下基于斯坦瑞树的最优资源存储方案，为了提供更好的服务通过将资源以斯坦瑞树的方式缓存在雾计算节点上，能使资源的存储达到最优化。Luan T H,Gao L,Li Z等在文献[5]研究了雾计算的主要关注点在于网络边缘的用户。作者详细阐述雾计算的发展，对云计算和雾计算在存储能力、计算能力和数据流通方面的对比。云和雾互相连接通信，云是布置在各处的雾服务器的中也控制点。每个雾服务器节点关注本地区的服务和云计算合作一起预测用户的需求。雾服务器会从云中选择不同的数据，然后存储到本地，并且实时的智能化更新本地存储的相关数据。不同位置的雾由中心云计算来统一协调和管理。Dimokas N,Katsaros D,Tassiulas L等在文献[16]中根据缓存节点的特点，提出了PCICC(Power Community Index Cooperative Cacheing)缓存策略。

目前大多数缓存决策策略大都没有综合考虑请求热点、网络能耗、内容流行度和雾服务器节点之间的协同等因素。雾服务器节点在缓存内容时容易忽略了各雾服务器节点之间的协同，导致整个缓存系统中出现缓存的内容冗余、内容种类的多样性较低以及节点中的内容被频繁替换的情况。

**二 雾计算中资源分配算法介绍**

缓存技术为了缓和计算机中央处理器和输入输出设备之间数据读写速度不匹配的问题，最早来源于计算机中的内存高速缓冲技术[17]。通过提前将处理器要访问的数据预先读取到高速缓存区域中，使得中央处理器能按自己的速度快速的获得需要的数据，从而迅速的处理任务，减少中央处理器中断的次数，提高中央处理器和输入输出设备的并行性能。随着网络的快速发展，高速缓存技术己经扩展到网络服务器应用的很多方面[18]。缓存从本质来说，就是采用缓存技术将要使用的数据提前复制后缓存到需要使用数据的区域。同理，在雾计算边缘服务器上提前缓存移动终端用户访问的数据，那么很多移动终端的数据请求便能够在本地的网络范围內快速获得响应，而不用每次都到云内容中心获取数据，能够实现少量外网带宽满足大量内网带宽需求的特点。

2.1缓存替换策略

缓存替换策略主要研究在缓存空间已满的情况下将哪个内容替换出去的问题。目前，缓存替换策略主要有LRU(Least Recently Used)、LFU(Leave FrequentlyUsed)、FIFO(First In First Out)和RAD(Random ReplacePolicy)。最近最少使用算法，即LRU。其核心思想是：如果某个数据最近被用户访问，那么该数据在将来被访问的概率也更高。最近最少使用算法通过Recency指标值的大小来决定是否需要置换掉缓存中存储最近最小使用的数据。当缓存空间不足的时候，LRU算法会把Recency值最大的数据置换掉，将新存入数据的Recency值设置为最小。如果用户访问的数据已经缓存在雾服务器中，那么系统会将该数据的Recency值设置为最小。最近最小使用算法复杂度较低，实现也较为简单，但是仅仅从最近的时间上考虑置换会影响整个缓存系统的性能，有时会将访问度高的数据置换出缓存中。LRU仅仅在特定的应用表现出较好的性[19]。近期最少使用算法，即LFU。其核心思想是：如果某个数据过去被访问过的频率高，那么该数据在将来被访问的频率也更高。LFU依据对数据的总体访问概率指标值的大小来判断是否需要置换掉缓存空间中的数据。当缓存空间不足的时候，通过LFU算法会把总体访问概率最少的数据对象优先置换掉。LRU算法的命中率比较高，但实现比LRU算法复杂。此外仅仅从最近访问频率上考虑置换数据对象，可能会置换掉某些缓存价值较高的数据。先进先出算法，即FIFO。其核心思想是：依据队列先进先出的性质，将最早存入的数最先被置换出缓存。先进先出算法的复杂度较低，实现的代价比较少，实现也比较容易，能够合理的利用缓存空间中存储数据的日志。但是该算法的命中率比较的低，因为最先存如缓存中的数据，很可能也是用户以后经常会访问的数据，因此在实际的开发引用中用的比较少[20]。 RAD随机缓存替换，其优点是不需要查找现有缓存中是否已经存在该内容。缓存替换策略在传统的Web缓存中已经得到广泛研究，详细内容可参考文献[21] 。由于本研究考虑的基于请求内容的缓存价值来决定替换，不同于传统的缓存替换策略，基于链路最小代价算法综合考虑了用户节点地理位置、请求内容流行度、雾服务器归属自治域内单位链路代价等因素。

这些常见的算法的特点一般来说是实现相对比较简单，但这些常见的算法的局限是没有办法针对应用变化自适应调节能力，而随着现有系统的不断升级，支持的应用的也在不断变化，因而对缓存算法的适应性也提出了更高的要求。

2.2缓存决定策略

缓存决定策略主要研究在某个雾服务器节点上是否缓存某个请求内容。在分布式架构中，可以通过云服务器来计算每个请求内容应该存放在具体的雾服务器节点的。依据协同决策点的不同，分布式缓存策略分为协作和非协作两种方式。本研究提出协作缓存策略正是基于雾服务器节点协作部分来缓存请求内容，从而降低对云内容中心服务器的负载，大大减少用户请求时延。即雾计算中缓存优化目标是在雾计算节点缓存空间大小限定的条件下，采用何种缓存优化策略，使得缓存的效益最优，能够提供最好的服务性能。在分布式缓存体系结构下，将缓存决定策略转化为最优问题进行求解，可以得到最优化缓存方案[22]。

2.3协作缓存策略

本研究所提出的基于雾计算中的协作缓存策略，即通过各个雾服务器节点对其自治域内用户提供内容缓存。每个雾服务器节点分为两部分：协作和非协作部分。自治域内的用户可以访问其所属雾服务器的协作和非协作部分，而非自治域内的其他用户只能访问该雾服务器的协作部分。通过控制协同比例（Ratio）来决定各个雾服务器对外提供的缓存大小，这样做的好处是非协同的部分放置自治域最热的内容，而协同部分则放置整个缓存系统中较流行的内容。同时通过这种协作缓存的方式可以大大减少缓存系统中内容冗余，提高缓存内容的多样性。

2.4协作缓存性能评价指标

雾服务器节点命中率除了受节点缓存容量的影响外还与缓存决定策略、替换策略和内容请求的概率分布等因素有关系。下面对影响命中率的相关因素介绍。

2.4.1内容流行度

内容流行度是指网络中内容被请求的概率。参考大多数文献，本研究也采用的Zipf[23]分布描述网络中内容的流行度，即网络中第（N为网络中内容块的总数目）个内容被请求的概率为：

 （2-1）

其中是Zipf参数，用于决定了内容的流行度分布。越小，内容的流行度分布约分散，越大，内容流行度分布越集中。通过研究表明，取值越大，对缓存命中率影响越大，反之亦然。

2.4.2 内容请求分布

参考大多数文献，本研究也采用泊松分布来描述用户发起的内容请求。假设内容请求服从参数为的泊松分布，则用户发出内容请求的时间间隔服从同参数的负指数分布，即在时间内用户发起两次内容请求的概率分布为：

 （2-2）

假设单位时间内内容的平均请求率为，则第个内容的平均请求速率为：

。 （2-3）

2.4.4 缓存命中率的计算

在给定的缓存系统中，雾服务器节点缓存容量一般是固定不变的。假设为，雾服务器节点采用LRU方式作为缓存替换策略，则在该雾服务器节点中对于任一内容的请求命中率为：

 （2-4）

请求命中率与缓存容量的关系为：

 （2-5）

其中N为网络中内容请求总数，为一个常数，其物理意义是连续两个内容请求命中的最大时间间隔。

对于给定，可以通过（2-5）求解出相应的[24]，继而可求得，以及缓存容量为时节点的缓存命中率：

 （2-6）

由此可见，在给定雾服务器节点缓存容量大小以及Zipf参数时，缓存命中率仅与当前网络中的内容块总量有关系。更确切地说，节点的缓存命中率仅与其缓存块中保存内容有关，所以本研究提出的链路最小代价结合实际网络请求内容，尽可能保存最流行内容，以达到提高缓存命中率。

2.5预测算法对缓存的影响

目前，研究人员已经开始研究主动缓存在无线网络中的应用问题。文献[11]讨论了通过预测技术预测人类动态行为的可能性，发现用户群体中93%的用户的移动行为能够得到准确预测。文献[25]利用概率图模型、人工神经网络、贝叶斯网络、决策树以及联合模型对用户的移动行为进行了预测，发现预测准确率达到60%，并且发现不同用户之间的行为差异化很大。文献[26]研究了移动网络中主动缓存方法的有效性，研究发现主动缓存算法相较随机算法能够提高85%的内容命中率，减少10%的流量开销。文献[27]讨论了协同过滤算法在内容热度预测工作上的应用，并讨论了机器学习算法在解决数据稀疏性方面的应用，同时根据智能移动设备之间可相互通信的特点，建立了社交网络模型，进一步讨论了主动缓存算法在提升缓存性能上的表现，研究[28]发现，主动缓存算法相较被动缓存算法能提升了26%的内容命中率，减少了22%的流量开销。以上研究论证了研究移动用户的行为习惯对改善网络性能的必要性，实验结果显示了主动缓存技术带来的性能优势。

**三 雾计算中资源分配工作展望**

本研究基于传统的云计算框架在支持移动智能终端设备方面的表现的不足，通过在数据中心和移动设备端之间拓展一层雾计算服务框架也就是引言中所提出的雾计算。雾层面的服务器通过非常近的距离与移动智能设备终端进行无线的相互连接，雾计算边缘节点缓存了大量的和其自治域相关的数据，实现了本地数据的存储优化和计算。雾计算框架中的雾计算边缘服务器的缓存能力和计算能力，可以更好的适应移动智能终端设备的需求特性，满足移动智能终端设备应用时延低的要求。然而雾计算节点存储空间有限，选择性的将数据存储在雾服务器节点是解决雾计算节点存储空间不足的问题、提升网络性能的好办法。本研究通过对雾计算边缘节点缓存进行分析，提出一种基于链路最小代价的协作缓存算法[29-36]。即通过对每个边缘雾服务器节点存储空间进行共享来达到完善整个缓存系统的存储空间。最后通过对各雾服务器节点的缓存记录进行分析，然后采用机器学习的方法来预测即将访问的内容，以达到更高的缓存利用效率。不同于传统被动式的存储[36-44]，主动式的缓存方式可以更大化满足不同自治域的变化需求，更加符合雾计算因地制宜、个性化定制的特点。

总之，随着移动互联网的不断发展，对资源的需求也在不断加大。依托云计算框架的强大计算能力和便捷性，为雾计算更加贴近用户终端提供发展的平台。雾计算是一个在网络资源优化领域尚属于全新的研究领域，随着计算的进步将会吸引越来越多的学者研究该领域，希望雾计算能加快推进产业化和实用化。