

# 一种基于 SDN 的 ICN 高效缓存机制

雷方元 蔡 君 罗建桢 戴青云 赵慧民

(广东技术师范学院电子与信息学院 广州 510665)

**摘 要** 内置缓存是信息中心网络(ICN)架构最核心的思想之一。为了充分、高效地利用全网缓存资源,提出了一种基于 SDN 的 ICN 高效缓存机制(SIC)。该方法利用 SDN 控制面感知全网状态,在 SDN 控制面中根据 ICN 缓存节点的网络社团重要度特性以及社团节点重要度特性来生成缓存策略,使不同流行度的内容对象在各社团之间和同一社团内节点上的分布更合理。在多种实验条件下对 SIC 策略进行了仿真验证,结果表明与 Hash+LRU、Betw+LRU 相比,该策略能更好地提升缓存命中率、跳数减少率、平均社团经过个数,同时 SDN 的资源开销保持在较低的水平。

**关键词** 软件定义网络,信息中心网络,缓存,替换,节点社团重要度

**中图分类号** TN915.9 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.8.015

## Efficient Caching Mechanism Based on Soft Defined Information-centric Networks

LEI Fang-yuan CAI Jun LUO Jian-zhen DAI Qing-yun ZHAO Hui-min

(School of Electronic and Information, Guangdong Polytechnic Normal University, Guangzhou 510665, China)

**Abstract** In-network caching is one of the most important features of ICN (information-centric networks). A caching strategy based on soft defined information-centric network was proposed to efficient utilize the whole network cache resource. In control plane, SDN controller aware the dynamic network status, the caching mechanism not only considers the importance of the cache community but also the node's importance within a cache community, which makes different popularity of content objects more reasonable in temporal distribution. The SIC mechanism was implemented under a variety of experimental conditions, and compared with the previous strategies Hash-LFU and Betw+LRU. The simulation results show that the mechanism can yield a significant performance improvement, such as, average request delay, cache hit ratio, hop reduction ratio, average community, and the resource overhead of SDN controller also keep the low level.

**Keywords** Soft defined network, Information-centric networks, Cache, Replacement policy, Node's importance to community

## 1 引言

据思科公司预计,到 2018 年,随着各种形式的多媒体业务的爆炸性增长,Internet 网络流量每年将增长 1.7 倍<sup>[1]</sup>。传统的 host-host 的 Internet 网络架构已经不能满足这种以用户驱动为主的内容获取方式。

研究者们提出众多新型的网络体系,信息中心网络(ICN)作为一种用户驱动的新型网络架构,在众多的网络体系中比较有优势。在 ICN 架构中,网络内置缓存是其最核心的思想之一,引起了研究者的广泛关注。目前的缓存方法分为两种类型:1)沿路(On-Path)缓存<sup>[8-13]</sup>,即数据包在返回过程中按照一定的策略选择部分沿途节点对该内容进行缓存,这种方式实现起来简单,能够充分利用沿路缓存资源,在一定程度上减轻了服务器的负载,但会浪费非沿路缓存节点

资源;2)旁路(Off-Path)缓存<sup>[15,17,18]</sup>,即数据包在返回过程的传输路径上按照一定的策略选择部分而不一定是沿路节点进行内容缓存,这种方式充分利用了旁路缓存资源,提高了缓存命中率,但在目前的 ICN 架构下实现起来较为困难。

软件定义网络(SDN)<sup>[2]</sup>是近年来备受关注的新型网络体系结构,其核心思想是控制与转发分离,两个面存在着双向交互的过程,在控制面,通过感知网络状态和网络资源,中央控制器对网络的逻辑控制功能和高层策略灵活地进行动态、优化配置;而在数据面,可在不影响网络正常流量的情况下执行这一配置,进而实现网络资源的优化配置。ICN 将主机地址与数据分离,使数据可以摆脱地址的束缚而存储于网络中任意位置。而网络的本质功能是要实现数据的网络供应与用户需求的匹配。SDN 的控制与转发分离的思想与 ICN 架构功能分离的设计原理天然相似,因而,最近 SDN 和 ICN

到稿日期:2015-05-19 返修日期:2015-07-30 本文受国家自然科学基金资助项目(61571141),广东省自然科学基金资助项目(2014A030310346,2015A030313672),广东省教育厅省级重大资助项目(2014KZDXM060)资助。

雷方元(1972-),男,博士,讲师,主要研究方向为 SDN、ICN, E-mail:leify@126.com;蔡君(1981-),男,博士,副教授,主要研究方向为未来网络、复杂网络;罗建桢(1984-),男,博士,讲师,主要研究方向为网络安全;戴青云(1965-),男,博士,教授,博士生导师,主要研究方向为图像自动识别;赵慧民(1966-),男,博士,教授,主要研究方向为压缩感知理论和信息安全技术。

融合的研究引起了研究者的广泛关注。为此,本文在 SDN 和 ICN 融合架构下,克服沿路缓存和旁路缓存的缺点,提出了基于 SDN 的 ICN 高效缓存机制 SIC,即以 Internet 网络拓扑结构社团为单位,从社团视角充分利用社团内的每一个节点的缓存资源,从而实现 ICN 的高效缓存。通过 SDN 控制面的控制策略,在网络内容被缓存时,不仅考虑每一个网络社团的位置,同时还考虑网络社团内部每一个节点的社团重要度,充分利用社团内部各个网络节点的缓存空间,将网络信息有效地缓存到相关节点。在节点缓存内容被替换时,通过 SDN 控制面,按照各个节点的重要性来指定替换内容;在节点缓存内容被访问时,通过 SDN 控制面,根据最优的路径将路由信息提供给访问者。从整体上达到网络缓存资源全局规划利用,优化网络流量,实现 ICN 高效缓存。

本文贡献如下:1)在 SDN 和 ICN 融合框架下,根据网络社团节点的重要度,利用 SDN 感知网络状态,提出了一种确定 ICN 缓存位置的决策策略,使内容对象在空间上分布更合理;2)基于 SDN 控制优化策略,根据社团重要度和内容对象在各节点处的访问情况,提出了一种 ICN 节点混合替换策略,使内容对象在时间上分布更合理;3)在多种实验条件下对 SDN 和 ICN 融合策略进行了仿真验证,结果表明与 Betw<sup>[11]</sup> 和 Hash<sup>[15]</sup> 相比,该策略能更好地提升包括缓存命中率、跳数减少率等在内的网络缓存性能指标。

## 2 相关研究

内置缓存是 ICN 架构的一个关键组件,其目标是使内容在网络中达到时间和空间上的合理分布,节约网络带宽,提升网络性能。从 ICN 数据包(Data Packet)响应兴趣包(Interest Packet)请求返回过程中选择缓存节点的方式可以分为:沿路缓存(On-path Cache)和旁路缓存(Off-path Cache)。

### 2.1 沿路缓存

沿路缓存即数据包在返回过程中按照一定的策略选择部分沿途节点对该内容进行缓存。在 ICN 架构原始提案中采用的处处缓存的机制(CEE)即将请求内容缓存于沿途传输路径上的所有节点中,该方法缓存空间利用率较低。Chai 等人<sup>[11]</sup>率先对该方法进行了改进,将请求内容缓存在沿途传输路径上介数最大的节点上,提出了一种基于节点介数的缓存方案(Betw),以节约缓存空间并提高缓存命中率。但是,在实际网络中,节点缓存空间远小于内容总量,这样将不仅导致流行度依然很高的内容被替换从而降低缓存命中率,而且会由于重要节点内容的频繁替换而导致网络整体性能下降。Psaras 等人<sup>[10]</sup>将数据包返回路径中的每个节点的缓存空间虚拟化为一个资源池,利用优化算法实现内容在沿路节点上的最优分布,即内容流行度高的内容优先存储于靠近内容源的节点处,该方法在层次化非常明显的网络拓扑结构中的缓存性能较优。Wang 等人<sup>[12]</sup>以降低沿路链路负荷为目标,采用启发式的方式来进行缓存节点的容量分配,实现最优的网络缓存节点容量分配。研究表明,网络拓扑结构和内容流行度是影响缓存容量分配的重要因素。

在沿路非协作的基础上,文献[13,14,19,20]提出在沿路缓存中进行协作式的方式来解决沿路节点之间缓存信息的交换,从而提高沿路节点的缓存效率。

Ren 等<sup>[13]</sup>将请求内容的流行度、缓存位置和缓存内容替换综合定义为局部缓存增益,并在兴趣包和数据包中增加字

段来携带增益信息。该方法中沿路中的每个缓存节点都需要计算并存储内容流行度,增加了缓存节点负荷。蔡君等<sup>[14]</sup>提出了以社团为单位,在数据包返回时将内容缓存到所经过的各社团内节点社团重要度最大的节点。在缓存内容替换时,根据节点社团重要度确定节点缓存队列中的内容对象被替换的概率,该方法适合于静态拓扑信息的社团网络。曾宇晶等<sup>[19]</sup>提出基于内容分块流行度和缓存节点位置的分级模型,其在用户侧的边缘路由器中增加兴趣统计 IST 表,该表用于统计用户内容请求数,并根据请求频率高低来对内容的流行度进行分级。该方法基于请求节点到缓存节点的跳数来确定沿路缓存节点的分级,并且仅仅在第一级缓存节点增加兴趣表,因此,适合于层次化结构网络。

沿路缓存方式的实现比较简单,能够充分利用沿路缓存节点的资源,提高缓存命中率。但是这类方法会大量浪费旁路缓存节点的缓存资源。

### 2.2 旁路缓存

旁路缓存(Off-path Cache)机制即内容在传输路径上按照一定的策略选择部分而不一定是沿路节点进行内容缓存。Saino 等人<sup>[15]</sup>在边缘路由器将用户请求的内容名称和数据包的名称通过 Hash 变换映射到某一核心路由器。这种方法利用了全局的缓存资源,但增大了边缘路由器的负荷,同时映射的 Hash 值存在冲突的可能。Shimizu 等人<sup>[17]</sup>将拓扑中的缓存节点进行编号,并把用户请求通过广播方式在树形拓扑的各分支节点进行传递。各个节点将用户请求与缓存中内容进行匹配,如果命中,缓存节点将其编号发送给请求节点,由请求节点确定最短下载路径的缓存节点。该方法为了避免网络风暴,采用树形拓扑结构。

Lee 等人<sup>[18]</sup>在兴趣包上附加扫描请求的方式获取到其它缓存节点的缓存内容。该方法中每个缓存节点的缓存过程完全独立,这会造成缓存内容冗余。葛国栋等<sup>[21]</sup>按照用户内容请求分布构建内容请求相似性社区 CSC,内容缓存时,将其转发到对应内容最活跃的区域中,并结合 Hash 策略在该区域中确定缓存节点。该方法中 CSC 的生成及 SAP 的动态交换都需要额外的系统开销。Wang 等人<sup>[22]</sup>提出基于缓存节点的中心度环模型的洪泛方式来获取指定节点周围 3 跳范围内的邻居节点的缓存内容,这种方式获得的缓存节点的增益主要适用于网络边缘的缓存节点。

目前,旁路缓存局部利用了非沿路缓存节点的资源,更进一步地提高了 ICN 的缓存效率。旁路缓存试图从全局的角度利用系统中的缓存资源,在一个分布式存储系统中,存储的位置的影响因素包括节点可用性、存储可用性、单个条目的内容流行度,因此,确定内容副本的最优位置是相当困难的,同时在目前的 ICN 架构下实现起来也很困难。

为此,本文在 SDN 与 ICN 融合框架下,利用 SDN 感知网络资源和网络状态,应用复杂网络社团思想,在全网范围内差异化地实现缓存放置和缓存替换,较好地解决了 ICN 缓存效率不高、控制不灵活等问题,真正实现了 ICN 的高效缓存。

## 3 基于 SDN 的 ICN 缓存策略

### 3.1 相关定义

定义 1(社团) 对 Internet 网络结构的实证研究显示,Internet 网络结构具有明显的社团特性,即在同一社团内部,节点之间的连接相对紧密,而在社团之间的节点之间的连接

相对稀疏;并且在同一社团内的大部分用户所关心的内容(兴趣)也是相似的。本文假设是在同一 SDN 控制器下,对不同社团的节点进行缓存管理。

**定义 2(节点社团重要度)** 基于节点在与不在网络中表示网络邻接矩阵的特征值的变化表示节点对网络社团特性的重要度,节点  $k$  的重要度  $P_k$  定义如下:

$$P_k = \frac{\sum_{i=1}^c v_{ik}^2}{v_i^2 v_i} \quad (1)$$

其中,  $c$  为网络中社团数目,  $v_i$  表示以网络中的路由器为节点、路由器之间的物理链路为边构建的邻接矩阵的第  $i$  个特征向量,  $v_{ik}$  表示特征向量  $v_i$  中的第  $k$  个元素。

为了归一化节点社团重要度,定义  $IN_k = P_k/c$ , 其满足  $\sum_{k=1}^n IN_k = 1$ 。

**定义 3(社团重要度)** 基于社团内节点社团重要度表示在同一控制器下不同社团的重要度,如:  $IC_k = \sum_{i=1}^{n_{C_k}} IN_i$ , 其中  $n_{C_k}$  表示社团  $C_k$  内的节点数目。

**定义 4(内容流行度)** 内容对象被访问的概率为内容流行度,以一个控制器下的节点或社团为单位,设其内容集合为  $CP = \{cp_i | i=1, 2, \dots, N\}$ ,  $CP$  中的内容  $cp_i$  按其流行度的级别高低进行标号,如对任意  $1 \leq i < j \leq N$ , 则被访问的概率  $p_i < p_j$ 。即流行度高的内容存储时,优先选择社团重要度大的社团,并进一步选择该社团中节点重要度大的节点,这些节点易于被其它的节点访问。

节点社团重要度用于在同一社团内,实现节点社团重要度与内容流行度的匹配,提高缓存空间的利用率。社团重要度用于在不同社团间实现社团重要度与内容流行度的匹配,更进一步提高缓存空间的利用率。

### 3.2 SIC 缓存策略机制

基于 SDN 的 ICN 缓存策略的实施框架如图 1 所示,控制面为 SDN 控制器,主要负责缓存策略的生成;数据面为 ICN 系统,除了实现正常的 ICN 路由等功能之外,还主要负责控制面生成的缓存策略的执行。控制面与数据面采用 SDN/Openflow 流表上报与下发形式进行通信。

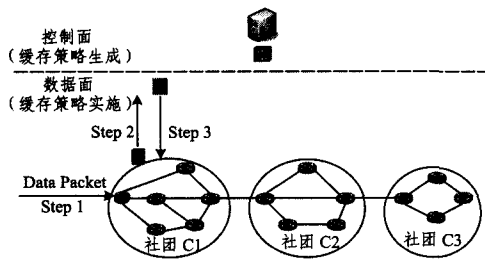


图 1 基于 SDN 的 ICN 缓存的实施框架

SDN 负责缓存策略的生成,包括确定缓存社团、确定缓存节点以及确定替换缓存的内容。前两步用于确定内容的存放位置,即缓存决策策略过程;最后一步用于确定替换缓存的内容,也即是缓存替换策略过程。

SDN 控制面中包含如下 3 张表:1) ICN 社团重要度表,主要用于记录各个社团的重要性;2) ICN 的社团结构表,用于记录 ICN 节点的社团节点信息、社团拓扑信息以及社团节点之间的路由信息;3) ICN 各个节点内容名称表,该表虚拟记录各个节点的内容存储的名称,主要用于快速检索缓存数据的名称。

下面简单说明 SDN 是如何利用这两张表来生成控制策略的。当 ICN 的节点  $A$  收到从数据源返回的数据包时,该节点将数据包的名称及内容流行度等信息通过流表形式上报到控制面。控制面收到数据包的名称后,根据内容流行度的大小,查找 ICN 社团重要度表来确定对应的社团;然后再检索对应社团中各个 ICN 节点内容名称表,来确定是否是首次经过对应的 SDN 控制器。如果是首次经过,则进一步根据数据包的内容流行度与 ICN 节点的重要度进行匹配,从而确定所需要存储的 ICN 缓存节点;如果缓存节点还有存储空间,则直接返回存储位置;否则要进行替换,进一步根据社团的重要度和节点重要度以及所经过的数据包的内容流行度计算替换节点的内容的概率。最后,SDN 控制器返回一个三元组  $(IC_i, IN_i, p_{ij}(k))$  给节点  $A$ ,  $IC_i$  表示需要缓存的社团,  $IN_i$  为需要缓存的节点,  $p_{ij}(k)$  表示对  $IN_i$  的缓存内容的第  $k$  个值以  $p_{ij}$  的概率进行替换。如果不需要进行缓存,则直接返回空值。数据平面收到信息之后,实施 SDN 的缓存策略。

这个过程充分利用了 SDN 强大的计算功能进行控制决策,使得整个 ICN 的各个节点的缓存资源得到充分、高效的调配和利用。

### 3.3 具体的实施过程

在基于 SDN 的 ICN 缓存策略实施过程中,SDN 作为上层控制面,它保存了 ICN 节点拓扑路由信息,同时动态更新 ICN 缓存节点的缓存以保持对 ICN 各个缓存节点的缓存情况均有详细的记录,所有缓存策略的生产计算过程均在 SDN 中完成,ICN 节点只负责缓存策略的执行。SDN 充分考虑网络的拓扑结构及每一个节点的缓存情况,从而使得整个 ICN 的缓存得到高效利用,避免了沿路缓存的资源浪费,同时降低了旁路缓存节点之间的通信开销。

SIC 的实施过程分为初始化阶段、缓存策略生成阶段和缓存策略执行阶段。SIC 算法整体的实施如图 1 所示。SDN 控制面完成缓存策略的生成,数据面负责缓存策略的实施。

#### 3.3.1 初始化阶段

(1) 对同一控制器下的缓存节点进行社团划分。

首先,依据社团重要度将社团从大到小进行排序。设一个控制器内的社团集合为  $IC = \{IC_i | i=1, 2, \dots, C\}$ ,  $C$  表示社团的个数,如对任意  $1 \leq i < j \leq C$ , 则  $IC_i < IC_j$ 。

(2) 社团内节点重要度的排序。

根据式(1)计算网络中每一个社团内各个节点的重要度,并进行归一化处理。

对于同一社团,首先根据节点社团重要度将节点降序排列,如对任意  $1 \leq i < j \leq N_{C_i}$ , 则  $IN_i < IN_j$ 。

#### 3.3.2 缓存策略生成阶段

在 SDN 控制器中执行缓存策略的生成,如图 2 所示,缓存策略包括确定缓存社团、确定缓存节点以及确定替换内容。

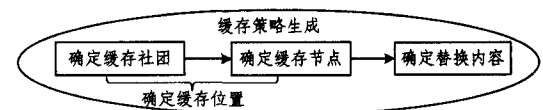


图 2 缓存策略生成机制

(1) 确定内容的放置位置,一是确定内容缓存社团的方法;二是在社团内确定内容缓存节点的方法。

确定内容缓存社团的方法:依据控制器下的内容流行度将内容均分为  $C$  个区间,同时比较需存储的内容的流行度与整个控制器下的内容的流行度,确定其所属的区间,从而确定

内容的缓存社团。

在社团内确定内容缓存节点方法:根据社团内的内容的流行度将内容划分为  $N_G$  个区间,同时比较需存储的内容的流行度与整个社团下的内容流行度,确定其所属的区间,从而确定内容的缓存节点。

### (2) 确定节点内的替换内容。

根据节点的节点社团重要度和内容对象的流行度,选择不同的内容对象进行替换。即通过节点社团重要度确定节点缓存队列中的内容对象被替换的概率,如节点社团重要度大的节点,在缓存队列中流行度低的内容对象被替换的概率大;而节点重要度小的节点,在缓存队列中流行度高的内容对象被替换的概率大,以更进一步地提高缓存空间的存储效率。若所选择的 ICN 节点的缓存空间大小为  $NC$  个对象,而该节点的节点社团重要度为  $IN_i$ ,该节点所属的社团的重要度为  $IC_j$ ,则该内容对象在该节点需缓存时,将按照概率  $p_{ij}(k)$  替换该节点缓存队列的第  $k(k \in [1, C])$  个位置的内容。 $p_{ij}(k)$  定义如下:

$$p_{ij}(k) = \frac{1}{\alpha} \left( \frac{IN_i}{IC_j} \right)^{\beta k} \quad (2)$$

其中,  $\alpha$  为归一化因子,满足式(3),  $\beta$  为概率调节系数。

$$\sum_{k=1}^C p_{ij}(k) = 1 \Rightarrow \alpha = \sum_{k=1}^C \left( \frac{IN_i}{IC_j} \right)^{\beta k} \quad (3)$$

### 3.3.3 缓存策略的实施阶段

将 SDN 控制面生成的缓存策略下发至交换机流表传至数据面,完成缓存策略的执行。

## 4 仿真实验与分析

### 4.1 SIC 实验模型

在仿真环境下利用 Openflow 流表来构建 SIC 实验模型控制面与数据面。ICN 的缓存节点网络拓扑采用 Xie 等人提出的 CBEN 模型<sup>[23]</sup>,该模型生成的网络具有明显的社团特性,用图  $G = \{G_1, G_2, \dots, G_Q\}$  来表示,其中  $G_i$  表示第  $i$  个社团。每个社团用  $G_i = (V_i, E_i)$  来表示,其中  $E_i$  表示第  $i$  个社团边的集合,  $V_i = \{v_{i1}, v_{i2}, \dots, v_{iN}\}$  表示第  $i$  个社团的节点集合。

网络中的内容用集合  $F = \{f_1, f_2, \dots, f_R\}$  表示,网络中内容服务器用集合  $S = \{s_1, s_2, \dots, s_p\}$  表示。假设网络中的每个内容随机分布在一个内容服务器  $s_i$  中,并且  $s_i$  仅与一个节点  $v \in V$  相对应;每个内容服务器的内容单元具有相同的大小,每个缓存节点的缓存空间大小相同,缓存存储器的一个缓存槽仅可容纳一个内容单元;内容请求到达服从  $\lambda = 50$  个/s 的泊松过程,用户对网络内容的请求符合 Zipf 分布,即用户对内容流行度为  $i(1 \leq i \leq M)$  的内容请求频率为  $P(X=i) = (i^{-\alpha})/C$ ,其中  $C = \sum_{j=1}^M j^{-\alpha}$ ,  $M$  表示内容的类别。

当用户发起内容请求时,SDN 控制器根据请求内容的内容流行度,在对应的网络社团中的对应节点中进行内容匹配,如果找到与之匹配的内容,则表示一次缓存命中;否则,表示一次缓存未命中。在缓存未命中的事件中,内容请求遍历整个内容分发路径直至内容服务器。在请求内容的数据包返回时,缓存节点将数据包的信息通过 SDN 控制器的 Openflow 流表同步到控制面,并将缓存决策  $(IC_i, IN_i, p_{ij}(k))$  通过流表下发至 ICN 缓存节点数据平面。

### 4.2 评价指标

基于 ICN 缓存线速执行的条件,提出如下要求:1)在有

限的缓存空间内缓存更多不同种类的内容;2)减少替换次数。SIC 研究的目的是充分利用网内缓存资源,降低网内缓存的冗余度,提高内容差异率。因此选取了资源开销、缓存命中率、平均经过社团个数、缓存路由跳数减少率作为评价标准,同时也特别考虑了内容流行度 Zipf 的参数对这些评价指标的影响。

在 SDN 与 ICN 融合框架下,ICN 缓存节点收到用户发出的兴趣包请求信息后,将信息通过 Openflow 流表发送到 SDN 控制器,由 SDN 控制器来定位出用户请求的位置,然后将相应的路由信息下发到请求节点,实现缓存查找。

为了衡量 SDN 与 ICN 的融合框架是否会由于 SDN 计算带来额外资源开销,引入资源开销来进行评估。资源开销 (Resource Overhead, RO) 定义为为用户发送请求兴趣包到接收到应答数据包之间 SDN 感知计算缓存资源的额外开销。

缓存命中率 (Cache Hit Ratio, CHR) 定义为在缓存中而不是由服务器端响应用户请求的概率,即网络中节点缓存内容响应兴趣包请求的概率。

平均经过社团个数 (Average Community, AC) 定义为为用户请求的数据包从最近缓存节点返回时所经过的社团个数。如果没有中间的缓存节点,AC 则表示请求节点与服务器节点之间的社团个数。随着中间缓存内容的增加,AC 会逐渐减小。同时,平均社团经过的个数反映用户感兴趣内容平均缓存的次数。

跳数减少率 (Hop Reduction Ratio, HRR) 表示由于网内缓存的存在,兴趣包命中时的跳数与兴趣包到服务器的跳数之比,其定义如式(4)所示:

$$HRR(t) = \frac{\sum_{r=1}^R h_r(t)}{\sum_{r=1}^R H_r(t)} \quad (4)$$

其中,  $H_r(t)$  和  $h_r(t)$  分别表示在  $[t, t+k]$  时间段内,用户从服务器和缓存位置获取内容  $f_r$  所需要的跳数。如果 ICN 网络中没有缓存,则  $h_r(t) = 0$ ,此时的 HRR 恒为 0。

### 4.3 实验结果

用于实验网络的拓扑由 100 个节点和 1004 条链路组成,其中描述网络的社团特性的模块度  $Q = 0.542$ ,  $Q > 0.3$  表明网络具有明显的社团特性。将本文的 SIC 方法与 Betw<sup>[11]</sup>、Hash<sup>[15]</sup> 和 Opportunistic<sup>[16]</sup> 进行对比,本文采用多个性能参数将 SIC 策略与 Betw+LRU 策略、Hash+LRU 策略和 Opportunistic 进行比较,其中包含系统的缓存命中率、跳数减少率、平均经过社团个数以及内容流行度的影响。其它参数的设计如下:用户数量设置为 5000,内容大小平均为 1MB,节点缓存也以 1MB 为一个单元。以下所有结果都是经过 10 次实验后得到的平均值。

图 3 给出了内容流行参数  $\alpha = 1.2$  和  $\alpha = 1.5$  时, Betw、Hash 与本文的 SIC 方案的缓存命中率 CHR 对比。随着  $\alpha$  取值的增加,内容请求的集中性和局域性不断加强,流行缓存资源的驻留概率和命中率不断增大,所有方案的 CHR 都得到了相应的提高。对于 Hash-LRU,域内所有节点整体协作,实现应答内容的协同缓存,虽无法实现内容的优化存储,但提升了域内缓存内容的命中率,缓存命中率提升幅度达 27.5%;对于 Betw-LRU,随着内容流请求更加集中,缓存命中率提升幅度达 29.1%。在 SIC 中,由于 SDN 的网络感知功能更进一步充

分利用了所有缓存节点的存储空间,增大了缓存内容的驻留概率和利用率,因此 CHR 分别达到了 61.4%和 70.3%。

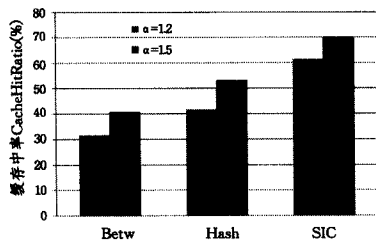


图3 内容流行度对缓存的影响

图4给出缓存空间大小对缓存命中率的影响,在各种方案中,缓存命中率均随着缓存空间增大而提高。在这些方案中,缓存空间增大后,可以容纳的缓存内容增多,用户发出请求时,缓存命中的概率也相应提高。随着缓存空间从10M增加到320M,CEE的缓存命中率从19.9%增加到32.5%;Betw方案的缓存命中率从25.9%增加到47.5%;Opportunistic从34.1%增加到53.4%;SIC的缓存命中率从61.4%增加到79.1%。

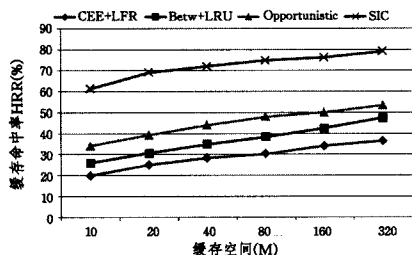


图4 缓存空间对缓存命中率的影响

图5给出Zipf分布中内容流行度参数 $\alpha$ 的值对缓存命中率的影响。随着内容流行度的增加,用户内容的请求的集聚度不断增加,在缓存空间一定的情况下,缓存中内容再次被请求的概率增加,缓存命中率得到提升。图5中,缓存流行度 $\alpha$ 从0.5增加1.8,缓存命中率从13.30%增加到89.98%。

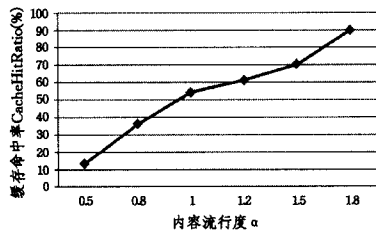


图5 内容流行度对缓存命中率的影响

图6为内容流行度对节点路由跳数减少率(HRR)的影响以及内容流行度对平均经过社团个数的影响。用户请求的数据包返回时,经过每一个社团都将根据其内容流行度存储在对应重要度节点的缓存空间中,因此,平均社团经过的个数反映用户感兴趣内容存储的次数,在存储时也会根据在对应节点是否有缓存空间而进行替换。在内容流行度 Zipf 参数 $\alpha=0.5$ 时,平均经过的社团为1.83个,平均节点路由跳数减少率 HRR 为 22.49%;随着内容流行度参数 $\alpha$ 的增加,平均经过的社团个数减少,跳数减少率 HRR 增加;当内容流行度参数 $\alpha=1.8$ 时,平均经过的社团个数减少到1.124个,平均的 HRR 为 35.43%。以上表明,随着内容集聚度增加,用户请求成功的概率和缓存效率越来越高。

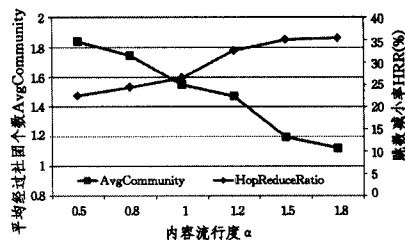


图6 内容流行度与节点路由跳数及通过社团数的关系

资源开销主要是 SDN 控制面的开销,分为存储开销和通信开销两个方面。

存储开销:为了进一步加快 SDN 控制器对节点内容名称表的查询速度,采用 MD5 哈希算法来对文件名称进行编码处理<sup>[25]</sup>。例如,假设内容名字采用 16 字节进行表示,内容流行度用 1 个字节表示,为了存储 100 万个内容名称,则存储需要的空间约为 16.212M。随着缓存价额的降低,在 SDN 控制器中为每个节点增加这样大小的内容名称空间是可以承受的。

计算开销:计算开销分为初始化阶段的社团重要性计算开销和兴趣包与数据包定位的开销两部分。

SIC 缓存决策策略中通过式(1)获得节点社团重要度,利用 Lanczos 算法和 QL 算法<sup>[24]</sup>求稀疏对称矩阵的所有特征值和特征向量的时间复杂度为  $O(lm)$ ,其中  $l$  和  $m$  分别表示网络的节点数和边数。这部分开销采用离线计算存储方式得到,因此可以忽略。

确定缓存对象所属的缓存节点的开销为  $O(nr)$ ,其中  $n$  为网络中社团的个数, $r$  为每一个社团内的缓存节点的个数。在一个缓存节点中利用 MD5 来确定缓存位置开销的时间复杂度为  $O(1)$ 。在替换过程中,计算概率的时间复杂度也仅为  $O(k)$ ,其中  $k$  为每个节点的缓存空间大小。

ICN 数据面通信开销:在 ICN 数据面中的兴趣分组和 Data 分组直接按照 SND 所确定的缓存策略来实施,不需要额外的通信开销来实施路径查找,因此,SIC 策略的通信开销很低,可以满足 ICN 对每个节点缓存处理线速的要求。

结束语 如何高效地利用 ICN 中每个节点的缓存功能,提高网络传输性能,成为了 ICN 架构应用于实际环境的关键问题之一。为此,本文在 SDN 与 ICN 融合的框架下,利用 SDN 控制面感知网络状态,在 SDN 控制面中根据感知的网络社团的重要性和社团网络节点重要度来确定缓存位置策略,并在此基础上生成缓存概率替换策略。在不同网络环境进行了仿真实验,结果验证了本文方法的有效性。后续的研究工作包括,网络拓扑的动态变化,即考虑网络节点的增删对缓存效率的影响。

## 参考文献

- [1] Cisco Visual Networking Index; Forecast and Methodology, 2014-2019[EB/OL]. <http://www.cisco.com>
- [2] Raghavan B, Casado M, Koponen T, et al. Software-defined Internet Architecture; Decoupling Architecture from Infrastructure[C]//Proceedings of the 11th ACM Workshop on Hot Topics in Networks. Redmond, Washington, USA, October 2012
- [3] Arumathurai M, Chen J, Monticelli E, et al. Exploiting ICN for flexible management of software-defined networks[C]//Proceedings of the 1st International Conference on Information-centric Networking. ACM, 2014; 107-116

(下转第 83 页)

- (PAM) Workshop, Boston USA, Springer, 2004; 158-167
- [10] Sarraute C, Burrone J. Using Neural Networks to Improve Classical Operation System Fingerprinting Techniques[J]. Electronic Journal of SADIO, 2008, 8(1): 35-47
- [11] Zhou Tie-zheng, Li Yuan, Zhang Bo-feng, et al. Operation system recognition based on support vector machines[J]. Journal of Tsinghua University (Science and Technology), 2009, 49 (S2): 2164-2168 (in Chinese)  
邹铁铮, 李渊, 张博峰, 等. 基于支持向量机的操作系统识别方法[J]. 清华大学学报(自然科学版), 2009, 49(S2): 2164-2168
- [12] Cheng Shu-bao, Hu Yong. Operating System Recognition based on Singular Value Decomposition and DAG\_SVMS[J]. China Information Security, 2013(9): 66-67 (in Chinese)  
程书宝, 胡勇. 基于奇异值分解和 DAG\_SVMS 的操作系统类型识别[J]. 信息安全与通信保密, 2013(9): 66-67
- [13] Shu G, Lee D. A formal methodology for network protocol fingerprinting[J]. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(11): 1813-1825
- [14] Kohno T, Broido A, Claffy K C. Remote physical device fingerprinting[J]. IEEE Transactions on Dependable and Secure Computing, 2005, 2(2): 93-108
- [15] Tom M. Mitchell. Machine Learning[M]. 增华军, 张银奎, 译. 北京: 机械工业出版社, 2013
- [16] Liu San-ming, Sun Zhi-xin, Liu Yu-xia. Research of P2P Traffic Identification Based on Decision Tree Ensemble[J]. Computer Science, 2011, 38(12): 26-29 (in Chinese)  
刘三民, 孙知信, 刘余霞. 基于决策树集成的 P2P 流量识别研究[J]. 计算机科学, 2011, 38(12): 26-29
- 
- (上接第 78 页)
- [4] Arumathurai M, Chen J, Maiti E, et al. Prototype of an ICN based approach for flexible service chaining in SDN[C]// 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs). 2015; 5-6
- [5] Gao S, Zeng Y, Luo H, et al. Scalable control plane for intra-domain communication in software defined information centric networking[J]. Future Generation Computer Systems, 2016, 56: 110-120
- [6] Liu J, Wang L, Zhang Y, et al. Hierarchical Caching Management for Software Defined Content Network Based on Node Value[M]// Advances in Parallel and Distributed Computing and Ubiquitous Services. Springer Singapore, 2016; 67-73
- [7] Cho K, Lee M, Park K, et al. Pack, Wave: Popularity-based and collaborative in-network caching for content-oriented networks [C]// IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs). 2012; 316-321
- [8] Arianfar S, Nikander P, Ott J. On content-centric router design and implications[C]// Proceedings of the Re-Architecting the Internet Workshop. ACM, 2010; 5
- [9] Psaras I, Chai W K, Pavlou G. Probabilistic in-network caching for information-centric networks[C]// Proceedings of the Second Edition of the ICN Workshop on Information-centric Networking. ACM, 2012; 55-60
- [10] Psaras I, Chai W K, Pavlou G. In-network cache management and resource allocation for information-centric networks [J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(11): 2920-2931
- [11] Chai W K, He D, Psaras I, et al. Cache "less for more" in information-centric networks (extended version) [J]. Computer Communications, 2013, 36(7): 758-770
- [12] Wang Y, Li Z, Tyson G, et al. Design and Evaluation of the Optimal Cache Allocation for Content-Centric Networking [J]. IEEE Transactions on Computers, 2016, 65(1): 95-107
- [13] Ren J, Qi W, Westphal C, et al. MAGIC: A distributed MAX-Gain In-network Caching strategy in information-centric networks[C]// 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs). IEEE, 2014; 470-475
- [14] Cai Jun, Yu Shun-zheng, Liu Wai-xi. Caching strategy based on node's importance to community in information-centric networks[J]. Journal on Communications, 2015, 36(6): 173-182 (in Chinese)  
蔡君, 余顺争, 刘外喜. 基于节点社团重要度的 ICN 缓存策略 [J]. 通信学报, 2015, 36(6): 173-182
- [15] Saino L, Psaras I, Pavlou G. Hash-routing schemes for information centric networking[C]// Proceedings of the 3rd ACM SIGCOMM workshop on Information-centric networking. ACM, 2013; 27-32
- [16] Hu X, Gong J. Opportunistic On-path Caching for Named Data Networking[J]. IEICE Transactions on Communications, 2014, 97(11): 2360-2367
- [17] Shimizu H, Asaeda H, Jibiki M, et al. Content hunting for in-network cache: Design and performance analysis [C]// 2014 IEEE International Conference on Communications (ICC). IEEE, 2014; 3172-3177
- [18] Lee M, Song J, Cho K, et al. Content discovery for information-centric networking[J]. Computer Networks, 2015, 83: 1-14
- [19] Zeng Yu-jing, Jin Ming-shuang, Luo Hong-bin. LICA: A Segment-Popularity Based Caching Scheme in ICN[J]. Acta Electronica Sinica, 2016, 44(2): 358-364 (in Chinese)  
曾宇晶, 靳明双, 罗洪斌. 基于内容分块流行度分级的信息中心网络缓存策略[J]. 电子学报, 2016, 44(2): 358-364
- [20] Wang J M, Zhang J, Bensaou B. Intra-AS cooperative caching for content-centric networks [C]// Proceedings of the 3rd ACM SIGCOMM Workshop on Information-centric Networking. ACM, 2013; 61-66
- [21] Ge Guo-dong, Guo Yun-fei, Liu Cai-xia, et al. Collaborative Caching and Routing Scheme Based on Local Request Similarity in Named Data Networking[J]. Journal of Electronics & Information Technology, 2015, 37(2): 435-442 (in Chinese)  
葛国栋, 郭云飞, 刘彩霞, 等. 命名数据网络中基于局部请求相似性的协作缓存路由机制[J]. 电子与信息学报, 2015, 37(2): 435-442
- [22] Wang L, Bayhan S, Ott J, et al. Pro-Diluvian: Understanding Scoped-Flooding for Content Discovery in Information-Centric Networking [C]// Proceedings of the 2nd International Conference on Information-Centric Networking. ACM, 2015; 9-18
- [23] Xie Z, Li X, Wang X. A new community-based evolving network model[J]. Physica A: Statistical Mechanics and its Applications, 2007, 384(2): 725-732
- [24] Newman M. Networks: an Introduction[M]. Oxford University Press, Inc, 2010
- [25] Bernardini C, Silverston T, Festor O. Mpc: Popularity-based caching strategy for content centric networks[C]// 2013 IEEE International Conference on Communications (ICC). IEEE, 2013; 3619-3623