基于后缀摘要的可选下一跳转发信息表聚合方法*

杜传震,田 铭,兰巨龙

(国家数字交换系统研究中心, 郑州 450002)

摘 要: NDN(named data networking) 网络直接依据层次化内容标志进行路由,内容条目数量激增使路由表规模呈现爆炸式增长。传统的单径路由方式对于转发信息表(forwarding information base, FIB)的聚合和压缩的作用已不大,为此提出一种基于后缀摘要的可选下一跳转发信息表聚合方法。一方面,将多可选下一跳的路由方式引入到转发信息表聚合过程,使得具有共同下一跳的内容条目进一步聚合,缩减了路由表项数量;另一方面,为解决前缀过度聚合带来的后缀空洞问题,利用布鲁姆过滤器提取后缀摘要,提高了路由的成功率。理论分析和仿真实验表明:将后缀摘要和可选下一跳 FIB 聚合相结合,可使路由表项缩减到原来的20%以下,同时随着布鲁姆过滤器哈希函数的增多,可使内容路由的成功率接近100%。

关键词:命名数据网络;转发信息表聚合;可选下一跳;后缀摘要

中图分类号: TP393 文献标志码: A 文章编号: 1001-3695(2014)01-0261-05

doi:10.3969/j.issn.1001-3695.2014.01.061

Nexthop-selectable aggregation based on suffix digest for forwarding information base

DU Chuan-zhen, TIAN Ming, LAN Ju-long

(National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China)

Abstract: NDN is a recently proposed general-purpose network architecture whose routing based on hierarchical content name. The explosion of the number of routing entries have driven the growth of routing table size at an alarming rate. The traditional methods of single-hop routing have little effects on FIB aggregation and reduction. This paper proposed a nexthop-selectable aggregation method based on suffix digest for FIB. On one hand, it put the nexthop-selectable method into the FIB aggregation to achieve a better aggregation with the same prefix and reduced the number of routing table size. On the other hand, it also provided the method based on suffix hole to make sure the success of routing. Theoretical analysis and results show that nexthop-selectable FIB aggregation method based on suffix digest can reduce the FIB sizes to below 20% and the success probability can be close to 100% with the growing number of hash functions of Bloom filter.

Key words: NDN; FIB aggregation; nexthop-selectable; suffix digest

0 引言

随着互联网技术与应用的飞速发展以及互联网用户的快速增长,传统 IP 网络中地址既表示节点位置信息又表示身份信息的方式混淆了位置和标志的功能界限,在支持内容分发业务上的局限性越来越明显。虽然类似于 CDN^[1]和 P2P 的增量式技术的提出和改进在一定程度上缓解了内容需求膨胀的压力,但其烟囱式、拼盘式改进和设计造成了网络复杂性高、可操作性差、全局优化困难、整体代价高等问题。

近年来,将内容与主机在网络层分离的改进方法引起了广泛关注。以内容为中心的网络 CCN(content-centric network)成为未来网络的一种重要模式和发展趋势。其中,NDN 提出一个全新的网络架构使用层次化的数据名字而不是 IP 地址进行数据传递,让数据本身成为因特网架构中的核心要素。针对NDN 网络的研究已经取得了部分成果^[2-5]。文献[6]提出了

NDN 网络架构中可扩展转发平面设计的概念、问题与原则,用于实现名字的快速查找及网络的大规模维护,但是并未给出详细具体的解决方案。文献[7]深入研究了 NDN 网络架构中自适应转发的问题,利用自带寻址信息的数据包来实现更高的可操作性及网络的恢复性,然而带来了技术上的可行性以及环境生存力的问题。文献[8]也提出了一种基于蚁群转发策略的NDN 邻居缓存路由方式,实现了邻居缓存节点的数据充分利用。现有研究解决了 NDN 网络架构中转发策略存在的问题,但在转发过程中的路由表项压缩及路由准确性等方面并没有给出具体的解决方案,NDN 网络架构中的单径路由方式在处理内容表项急剧膨胀带来的问题上仍具有局限性。文献[9]提出的一种可选下一跳 FIB 聚合有效地解决了 IP 网络路由表剧增问题,但并不适用于命名数据网络。本文在其基础上作进一步研究,引入动态规划的思想对 NDN 网络原有的聚合方式进行改进,有效地压缩了路由表项。然而前缀过度聚合会导致

收稿日期: 2013-04-09; 修回日期: 2013-05-23 基金项目: 国家"973"计划基金资助项目(2012CB315901);国家"863"计划基金资助项目(2011AA01A103)

作者简介:杜传震(1990-),男,山东滕州人,硕士研究生,主要研究方向为宽带信息网络、新型网络(duchzhen@126.com);田铭(1984-),女,河北辛集人,硕士研究生,主要研究方向为宽带信息网络、新型网络;兰巨龙(1962-),男,河北张北人,教授,博士,主要研究方向为高速路由器关键技术、新型网络体系结构.

后缀空洞问题^[3],即只有部分内容条目的内容路由器聚合后却发布"基于该前缀下的内容对象都可通过我获取"的消息,这种现象会导致信息的大量丢失。针对这个问题,利用布鲁姆过滤器集合查询的特性^[10],提取相应条目的后缀摘要以确保路由的准确性。

本文介绍了 NDN 网络架构的命名方式及其路由整体框架,引入多径路由的思想,在命名数据网络中实现可选下一跳 FIB 聚合,针对大量前缀聚合引起的后缀空洞问题,提出基于布鲁姆过滤器后缀摘要方法,并对提出的聚合方法进行实验仿真。

1 NDN 网络架构模型

NDN 网络设计方案中直接采用层次化的数据名字(content object name, CON)进行路由,其命名机制类似于目前的统一资源定位符(URL)命名方案^[11-13]。例如"/example. com/videos/HarryPotter. mpg"为一个具体内容的名字(name),而"/example. com"及"/example. com/videos"则可作为内容前缀(prefix)用于路由查找及转发。层次化的名字可以实现分级聚合且易于表示信息的集合,在扩展性上具有很大优势。

NDN 网络的路由方式与 IP 相似, NDN 路由器发布名字前缀公告并通过路由协议在网络中传播,每个接收到公告的路由器建立自己的转发信息表(FIB)。传统的路由协议如 OSPF、IS-IS、BGP 都适用于基于名字前缀的路由。当有多个兴趣包同时请求相同数据时,路由器只会转发收到的第一个兴趣包,并将这些请求存储在未决请求表(pending interest table, PIT)中。当数据包传回时,路由器会在 PIT 中找到与之匹配的条目,并根据条目中显示的接口列表分别向这些接口转发数据包。转发完成后,路由会删除相应的 PIT 条目,并将数据包储存在内容存储库(content store, CS)中,如图 1 所示。



图 1 内容路由器转发过程

NDN 网络中采用转发信息库 FIB 聚合^[14]具有独特的优势,可以实现路由器级别的增量部署且不影响转发,如图 2 所示。然而,单一下一跳 FIB 聚合只能减少部分 FIB 条目,并不能缓解大量内容流量导致的路由表急剧膨胀带来的压力。

FID	
前缀	接口列表
/examples	0,1

图 2 FIB 转发信息表结构

2 可选下一跳 FIB 聚合描述与方法

相对于最短路径优先的单一下一跳路由,本文采用多可选下一跳路由的思想,通过在同一源和目的节点间找出多条路径同时参与数据传输,提出可选下一跳 FIB 聚合方法(nexthop-selectable FIB,NS-FIB)。该方法最大的特点就是针对每个聚合的前缀都有对应多个可选下一跳,使得具有公共路径的前缀进行聚合用于提高路由表项的压缩率。

2.1 基本思想

在 NDN 网络架构中,FIB 前缀聚合只需满足两个条件:a) 层次化命名结构中,转发的内容名字具有共同的前缀;b)它们 具有共同的下一跳。

举例说明,首先假设两个路由条目: $\langle \text{/examples/video}, a \rangle$ 和 $\langle \text{/examples/movie}, b \rangle$,它们没有共同的下一跳,因此无法聚合。再假设下一跳 $a \lor b$ 都可以将数据包"/examples/video"传送到目的地; $b \lor c$ 都可以将/examples/movie 发送到目的地。取代针对每一个前缀只有单一下一跳的方式,采用选择两个路由条目 $\langle \text{/examples/video}, \{a,b\} \rangle$ 、 $\langle \text{/examples/movie}, \{b,c\} \rangle$ 。这两个数据名就可以聚合为 $\langle \text{/examples}, b \rangle$ 。

NDN 网络架构中,针对层次化的命名方式,可选下一跳FIB 仍然是一个二元组集合: $F = \{\langle p, A_p \rangle \mid p \in P \}$,其中 P 为NS-FIB 的前缀集合,而 A_p 为前缀 p 的可选下一跳集合。针对该聚合同样为一个二元组集合: $F_{\text{aggr}} = \{\langle p', a_{p'} \rangle \mid p' \in P' \text{ and } P' \subseteq P \}$ 。 F_{aggr} 是一个可行的聚合,当且仅当 $\forall p \in P$ 和 p 在 F_{aggr} 中的最长匹配前缀为 $p', a_{p'} \in A_p$ 。 也就是说,对任意前缀,在 F_{aggr} 中得到的下一跳是 F 中的可选下一跳。可选下一跳 FIB 聚合问题:给定任意聚合单元 F,找到最优可行聚合 F_{aggr} ,使得 $|F_{\text{aggr}}|$ 最小。

首先采用压缩的多叉树来存储 FIB,如图 3 所示。给出了上述可选下一跳 FIB 的属性结构 $^{[15]}$,其中每个根节点都对应该子树的共同前缀。如果某个子树的每个节点拥有某个公共下一跳,那么该子树可以聚合到其根节点对应的前缀。 p_0 、 p_1 、 p_2 、 p_3 和 p_8 都拥有可选下一跳 a,它们可以聚合为 $\langle p_0, a \rangle$ 。称这样的子树为一个聚合单元。NDN 网络运用可选下一跳 FIB聚合算法可以找到互不相交的聚合单元进行最优聚合,利用动态规划计算出对应公共前缀的长度。

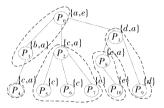


图 3 树形式的可选下一跳 FIB

2.2 聚合算法

定义 G(T) 为聚合单元 T 的最优聚合大小。 $G_x(T)$ 表示聚合单元 T 以 x 为下一跳的最优聚合大小,显然

$$G(T) \leq G_x(T) \rightarrow G(T) = \min_{x \in A_{R_x}} G_x(T)$$
 (1)

一个树型结构聚合单元 T 包括一个根节点和它所连接的所有子孙。 S_p 表示以 p 为根节点的一个分支,即包括节点 p 和其所连接的所有子孙。 C_r 则表示聚合的父亲节点集合 T的所有孩子节点。 $U_x(p)$ 表示以 p 为根、以 x 为下一跳的整个聚合树集合,其中 $x \in A_p$ 。

将可选下一跳 FIB 聚合单元 T 分解为一个以 R_T 为根、以 x 为下一跳的聚合集合 T 和以 T 的孩子节点为根的分支集合。 T 可以被聚合为一个条目,在此基础上的最优聚合大小为 $1+\sum_{p\in C_T}G(S_p)$ 。 T 可以是任意形式的组合,只要它是以 R_T 为根以 x 为下一跳的集合。其中, $G_x(T)$ 取所有不同形式中最小值。

$$G_x(T) = 1 + \min_{T' \in U_x(R_T)} \sum_{p \in C_{T'}} G(S_p)$$
 (2)

例如,假设 x=a,一种可能性聚合为 $T'=(p_0,p_1)$, $S_{p_4}=(P_4)$, $S_{p_2}=(P_2,P_5,P_6,P_7)$ 和 $S_{p_3}=(P_3,P_8,P_9,P_{10})$ ($G(S_{p_3})=2$)。

另外一种可能性为 $T' = (p_0, p_1, p_2, p_3, p_8)$, $S_{p_4} = (P_4)$, $S_{p_5} = (P_5)$, $S_{p_6} = (P_6)$, $S_{p_7} = (P_7)$, $S_{p_9} = (P_9)$, $S_{p_{10}} = (P_{10})$ 。可以发现, 基于第一种方法的聚合大小为 4,第二种方法的聚合大小为 6。实际上,第一种方法获得了以 a 为下一跳、以 R_T 为共同前缀的最优聚合,如图 4 所示。

$$\begin{split} & \text{Aggr. 1} \rightarrow \{ \langle p_0, a \rangle \langle p_2, c \rangle \langle p_3, d \rangle \langle p_4, c \rangle \langle p_8, e \rangle \} \\ & \text{Aggr. 2} \rightarrow \{ \langle p_0, a \rangle \langle p_4, c \rangle \langle p_5, c \rangle \langle p_{10}, e \rangle \langle p_6, c \rangle \langle p_7, c \rangle \langle p_9, d \rangle \} \end{split}$$

图 4 两种聚合结果比较

用特殊值初始化 G(T) 和 $G_x(T)$:

当|T|=1 时,G(T)=1 表示单一节点树的最优聚合大小为 1。 当 $x \notin A_{R_T}$ 时, $G_x(T)=\infty$ 表示选择一个不可用的下一跳是不允许的。

考虑到 $|U_x(R_T)|$ 的大小是指数级的,会造成算法的复杂度增加。采用动态规划的思想首先要考虑一个聚合树中的一个分支 T,让 T_x^* 来代表以 R_T 为根、以 x 为下一跳的最优聚合小单元,使其满足 $\sum_{p\in C_{T_x^*}}G(S_p)=G_x(T)-1$ 。 Aggre-cell()算法可以得到以 x 为下一跳、以 R_T 为根的最优聚合小单元 T_x^* 及 $G_x(T)$,用于简化查询 $U_x(R_T)$ 的过程。 Aggre-cell()算法能够估算 R_T 的每个孩子节点,将以孩子节点为根的聚合单元连接在一起。因此,Aggre-cell()算法的复杂度为 $\theta \mid C_{(R_T)} \mid$,仅仅与孩子节点的数目有关系。

算法 1 以 x 为下一跳的聚合单元 T 聚合算法 Aggre-cell (x,T)

条件: $x \in A_{Rr}$

- a) 初始化聚合单元 $T' = \{R_T\}$, $G_p = \infty$ 。
- b)对 R_T 的孩子节点 p 逐个查询, 判断是否满足 $x \in A_p$, 且 $G_x(S_p) = G(S_p)$, 若满足, 执行步骤 e); 否则, 继续下一查询。 若都不满足,则执行步骤 e)。
- c)若 $G_p > \sum_{p' \in C_{R_T} \& p' \neq p} G(S_{p'})$, 则令 $G_p = \sum_{p' \in C_{R_T} \& p' \neq p} G(S_{p'})$ 且 $p^* = p$; 否则,继续下一查询。
- d) 结束循环,令 $T' = T' \cup (S_{p^*})_x^* \mid (S_p)_x^* \neq U$ 是以 p^* 为根的最优化聚合单元 \ 。
- e) \diamondsuit T_{x}^{*} = T' , $G_{x}(T)$ = 1 + $\sum_{p \,\in\, C_{T'}} G(S_{p})$ $_{\circ}$
- f)结束算法。

算法 1 针对以 P_T 为公共前缀的聚合单元,搜索对应孩子节点的所有可选下一跳的最优聚合方案。将查询的孩子节点与根 P_T 结合,形成关于 T 的所有以 x 为下一跳的最优聚合。

算法 2 聚合单元 T 最优算法 Aggre-NS-FIB(T)

- a) 对于所有 $T \in V_T$, 初始化 $G(T) = \infty$ 。
- b)对于所有的可选下一跳 $x \in A_{R_T}$,利用 Aggre-cell (x,T) 计算得到 T_x^* 与 $G_x(T)$,如果 $G(T) > G_x(T)$,执行步骤 c);否则,继续下一循环。 c)令 $G(T) = G_x(T)$,当前 x 作为 R_T 的最优可选下一跳, R_T 的直
- 接子类作为 $C_{(S_p)_{x^*}}$ 。 d) 结束算法。

算法 2 是将 Aggre-cell(x,T) 得到的所有聚合方式进行迭代比较,得出具有最优可行性聚合单元(主要是最优下一跳选择与最大聚合数的折中考虑)。

2.3 构造方式

NDN 网络采用 LFA 方法^[16] 构造多条可选路径即多个下一跳,因为 LFA 和域内路由协议 OSPF(open shortest path first) 具有很好的兼容性,适合于内容网络 NDN 的数据传输。LFA 机制中存在两个条件来选择备选路径 LFC(loop free condition) 和 DSC(downstream condition)。

- 1) LFC 对路由器 R 而言,如果一个邻居 N 到目的地址不经过 R,那么 N 是一个 LFC 下一跳。如果多个路由器同时采用 LFC 条件构造可选下一跳 FIB,那么可能会出现转发环路。因此如果要大规模部署 NS-FIB 聚合方案,需采用另一种方法。
- 2) DSC 对路由器 R 而言,如果一个邻居 N 到目的地址的距离小于 R,那么 N 是一个 DSC 下一跳。由于这个条件满足了单调性原则,因此即使大量路由器部署,该聚合方法也可以保证转发无环。

3 后缀摘要提取方法

3.1 后缀空洞与布鲁姆过滤器

NDN 网络中采用的可选下一跳 FIB 聚合方法虽然提高了 FIB 表的压缩比率,缓解了路由表膨胀带来的压力和问题,然 而过多过分散的聚合会导致后缀空洞现象的发生。

NDN 网络架构中,如果所有具有相同前缀的内容名字存储在某个路由节点,只需要广播一个路由通告,其他路由器也只需要保存、维持一个条目就可以将相应的请求转发到这个节点上^[17]。然而,内容对象总是缓存或复制在多个位置,这使得基于前缀的汇聚有效性逐渐降低,即网络中的缓存节点通常不会存储到具有相同前缀的所有内容对象。如果基于前缀的聚合用在上述节点中来解决路由表项膨胀问题,由于路由通告仅能表达"基于该前缀下的内容对象都可通过我获取",将会导致错误的路由判断且丢失大量信息,称为后缀空洞现象(suffix hole)。该问题导致了内容定位时的不确定性,降低了路由效率。在最坏的情况下,路由器需要将匹配某一前缀的请求发送给所有匹配该前缀的节点,得到多个端口的多个响应。标准布鲁姆过滤器原理如图 5 所示。

本节主要针对后缀空洞问题,利用布鲁姆过滤器可以实现集合查询,常数时间和存储空间开销小、可以存储内容摘要的特性(结构如图 6 所示),对 NS-FIB 聚合加以改进。提出基于后缀摘要的聚合方法:在聚合单元内容名字的层次化结构中,利用布鲁姆过滤器得到每层名称相应的哈希值,确保路由判断的正确性,解决了路由通告与查询的准确性问题。

3.2 后缀摘要提取算法

为方便设计和观察,假设 NDN 提出的层次化的名字具有足够长度的 k 层结构(如若长度不够 k 层的,可以认为名字之后的层级为零或不存在)。在一个聚合单元中,最优聚合的匹配前缀长度为 m 层,则一个条目所剩下的后缀长度为 k-m 层,如图 6 所示。

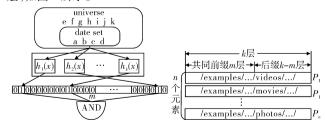


图 5 标准布鲁姆过滤器原理

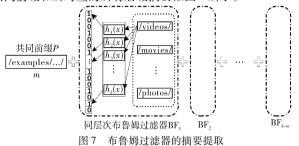
图 6 聚合元素的分层结构

有两种构造方式可以选择,即单一布鲁姆过滤器查询和层级布鲁姆过滤器查询。

单一布鲁姆过滤查询算法的应用原理是将对应每一个聚合元素的后缀 k-m 层作为一个元素进行哈希映射,这种方式会由

于内容膨胀和名字数量剧增而造成大量的假阳性误判,即布鲁姆 过滤器中某元素不属于数据集合而被指称属于该数据集合。

采用层级布鲁姆过滤器查询算法,对聚合单元相应的同层级不同的名称哈希生成后缀摘要。每层级对应一个布鲁姆过滤器来进行存储摘要,每个布鲁姆过滤器的元素集为该层级的不同元素的名称,其中\P\为所聚合的元素集。聚合后的条目形式为"/P/digest₁/···/digest_{k-m}"内容路由器通告聚合单元的共同前缀和所对应的所有后缀摘要如图 7 所示。



对于 N 个 CON (前缀为"/example. com/video"),一个像 "/example. com/video/digest₁ (titles)/digest₂ (categories)/…"形式的 sCON 可以用来代表这些内容对象的名字,其中 digest 的数目为 k-2 个。一个 sCON"/example. com/digest (categories/titles)"可以用来代表具有同前缀"/example. com/"但不同类别和标题的内容对象名字 CON。NDN 网络采用布鲁姆过滤器后 FIB 转发表的结构如图 8 所示。



图 8 改进后的 FIB 结构

在 NDN 网络的信息转发表聚合过程中,首先运用可选下一跳 FIB 聚合算法得到所聚合元素集合以及它们的共同前缀 P_r 的长度,而后从前缀之后的第一个层级到最后一个层级 $\mathfrak{A}^{[18]}$,采用布鲁姆过滤器提取后缀摘要。

算法 3 后缀摘要算法 suffix-digest(T)

- a)利用 Aggre-NS-FIB(T)得出聚合单元的共同前缀 P_T 。
- b) 查询得到前缀 P_T 的层级数目 m_o
- c)对于所有聚合元素的后缀,依次查询i层级名字并进行哈希处理存储到相应的布鲁姆过滤器 BF_i 中,如果名字不存在,则设为0。

d)结束算法。

为询问一个相应的 CON 是否存在,被询问的 CON 应该和一个 sCON 相匹配,即相应的汇聚前缀应该一样,并且在聚合单元 sCON 中应该给出一个匹配响应来表明在被询问的 sCON 中相应的后缀元素是可以表示出来的。应当注意,如果用布鲁姆过滤器算法的方法会发生假阳性误判,造成后缀摘要之后会错误地表明这个元素在本摘要中。降低查询误判对应用的影响需要通过设计合适的过滤器(如计数式布鲁姆过滤器(counting Bloom filter)查询算法、压缩布鲁姆过滤器(compressed Bloom filter)查询算法等)、增加哈希函数的数量 w 或增大向量空间的大小 v 等方法。一个发布者通过过滤器可以灵活地控制基于内容流行度的聚合水平,有效地维持路由状态所需要的网络资源中内容定位之间的平衡以及误报率。

3.3 参数选择与复杂度分析

布鲁姆过滤器作为一种支持集合查询的数据结构,在其实

现高效、简洁地表示集合效果的同时,却存在某元素不属于数据集合而被指称属于该数据集合的可能性,即假阳性误判。因此,在实际应用中必须对查询误判率进行评估,设计合适的布鲁姆过滤器,降低查询误判对应用的影响。

首先用 p 来表示向量空间中任一 bit 位为 0 的概率,那么任一 bit 位为 1 的概率为 p-1。假设哈希函数取值服从均匀分布,则当集合中所有元素都映射完毕后,向量空间中任意一位为 0 的概率为

$$p' = (1 - \frac{1}{v})^{wn} \approx e^{-\frac{wn}{m}} = p$$
 (3)

当不属于集合的元素而误判断属于集合时,元素在向量空间的对应位置都必须为1。元素的误判率为

$$f^{\text{BF}}(v, w, n) = (1 - p')^{w} \approx (1 - p)^{w}$$
 (4)

即 $f^{BF}(v,w,n) = (1-p)^w = (1-e^{-un/v})^w = \exp(w \ln(1-e^{-un/v}))$ (5) 由公式推知, $w = w_{\min} = (\ln 2)(v/n)$ 时,可以使 $w \ln(1-e^{-un/v})$

$$p = (1 - p) = \frac{1}{2} \tag{6}$$

即为在设计布鲁姆过滤器的时候考虑为最小误判率时,布鲁姆过滤器向量这种由 0 和 1 组成的随机位串中任意位置为 0、1 的概率均为 1/2。那么元素出现查询误判率最小值为

$$f(w_{\min}) = (1/2)^{w_{\min}} = (1/2)^{(\ln 2)^{\frac{v}{n}}} = (0.6185)^{\frac{v}{n}}$$
 (7)

w 是哈希函数的个数,应为整数,即

e-wn/v)达到最小值,这时

$$w = \lceil \ln 2(v/n) \rceil \tag{8}$$

式(8)给出了聚合单元中聚合元素个数n后,如何选择向量空间v与合适的哈希函数个数w使得布鲁姆过滤器误判率f最小。可以通过向量空间与哈希函数个数的关系来选择合适的参数,使得误判率f最小。

在 NDN 网络 FIB 聚合机制中使用标准布鲁姆过滤器,增加或判断一个层级名称元素都只需要进行 w 次哈希运算,其操作的时间复杂度为 O(w)。对于同一层级 n 个名称元素的集合,只需要 v 位向量空间,其空间复杂度为 O(v)。使用布鲁姆过滤器完成集合存储,只需要为每个元素平均保存 v/n 位,十分简洁。布鲁姆过滤器查询算法所需空间与元素自身大小无关,仅与元素映射到向量的位数相关,因此对 FIB 转发信息表的空间大小影响很小。

4 仿真实验及结果

4.1 实验设置

采用 C++及 MATLAB 语言模拟一个 NDN 网络运行过程 拓扑图并进行相关性能仿真。根据动态规划思想实现路由节点 FIB 聚合结构,节点对收到的数据包名字进行 FIB 聚合。由 BRITE 拓扑生成工具生成一个路由节点数为 100~500 的平面随机网络拓扑,其中任意两个路由节点之间存在直接相连接路径的概率为 0.2~0.5。从边缘节点中随机选取一个节点作为服务器节点,剩下的节点作为普通的内容路由节点,具备 CS表与 FIB表。为了便于分析,相邻节点之间的延迟设为 10 ms。

内容路由节点与用户主机直接相连,由用户发送兴趣包(数据请求包)。在模拟用户发送数据包时使其满足泊松到达特性^[19](λ = 200 个/秒)。

内容服务器共有 20 000 个内容对象, 假设采用 7 层级 URL 结构的名字作为内容在网络中的标志, 每层级的名称随

机为五种不同元素,这样不相交的内容名字共有 5⁷ 种。抽取 其中的 20 000 个作为内容名字与内容序号 1~20 000 ——对 应,用户请求的内容与其相应。

设定路由节点具备 NDN 路由结构模型,最多存储 400 个内容,FIB 大小设为 500。为了方便对比 LFC 聚合以及 DSC 聚合方式,对聚合前后 FIB 的大小进行比较,将 FIB 聚合压缩率作为衡量标准,即压缩后与压缩前的 FIB 大小比值。CS 表的缓存替代策略为 LRU 替换。

4.2 实验结果

4.2.1 NS-FIB 聚合分析(图 9)

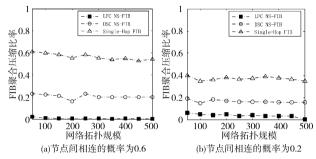


图 9 FIB 聚合压缩率随着网络拓扑大小的变化

从图 9(a)可以看到,采用单路径 FIB 聚合的 NDN 网络路由表项只能压缩到 60%。如果采用 LFC 可选下一跳 FIB 聚合方法,FIB 聚合可压缩至 0.23% ~ 2.43%,甚至用 DSC 方法构造,路由表项大小也可以缩减到 20%以下。

图 9(b)降低了路由节点之间连接的概率,仍然能得到同样的聚合效果。单路径聚合的效果变好是因为邻居节点的减少提高了前缀都指向相同下一跳的可能性。

路由节点连接度对 FIB 聚合压缩比率的影响如图 10 所示。随着概率的提高,连接度变大,单路径 FIB 聚合的压缩效果逐渐降低,这是因为邻居节点增多导致具有共同下一跳的前缀减少了。但 NS-FIB 聚合方法影响不大是因为邻居节点变多导致可选择的下一跳也增多了。

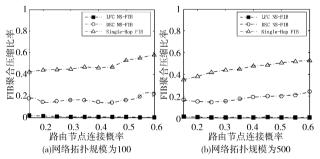


图 10 FIB 聚合压缩率随着路由节点连接度的变化

4.2.2 后缀摘要结果分析

后缀空洞现象在实验中具体表现为节点中不存在对应兴趣包的数据包,随着具有共同前缀的内容增多,节点能存储到全部对应的内容名字概率越小,即请求的命中率下降,如图11所示。

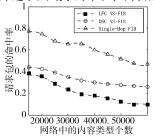


图 11 FIB 聚合中的后缀空洞现象

对基于后缀摘要的 FIB 聚合进行模拟仿真,考虑到内容数量对请求包命中率的影响,选用不同数量的内容对聚合进行实验。如图 12 所示,随着哈希函数个数增多,后缀空洞现象会逐渐消失,内容路由的成功率会越来越接近 100%。将 NS-FIB 聚合方法与布鲁姆过滤器相结合,有效地改善了路由环境,提高了内容路由的效率,解决了路由表项膨胀的问题。

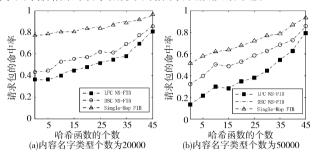


图 12 FIB 聚合请求包的命中率与哈希函数个数的关系

5 结束语

从优化聚合 NDN 网络架构转发信息表的角度出发,本文首先利用动态规划的思想提出一种多可选下一跳 FIB 聚合方法,缩减了路由表项;其次又提出一种基于布鲁姆过滤器的后缀摘要方法解决过度聚合造成的后缀空洞问题,提高了路由的准确性。理论分析和仿真实验都表明,将后缀摘要提取方法与NS-FIB 聚合方法相结合,能够实现有效地压缩 FIB 转发表和消除后缀空洞现象的作用。然而研究中发现 NDN 网络内容路由表中给出的只是到服务器端永久副本的下一跳路径。如何构造快捷路由思想,即能够建立到可能暂时缓存内容的未知近邻路由器节点路径的方法,将是下一步研究的重点。

参考文献:

- PALLIS G, VAKALI A. Insight and perspectives for content delivery networks [J]. Communications of the ACM, 2006, 49(1): 101-106.
- [2] ZHANG Li-xia, JACOBSON V, TSUDIK G, et al. Named data networking(NDN) project[R/OL]. (2011-05-20). http://named-data. org.
- [3] JACOBSON V, SMETTERS D K, THRONTON J D, et al. Network named content [C]//Proc of the 5th International Conference on Emerging Networking Experiments and Technologies. 2009:1-12.
- [4] JACOBSON V, THORNTON J, PLASS M, et al. Project CCNx[™] [EB/OL]. (2012-03). http://www.ccnx.org.
- [5] 唐晖, 周旭, 韩言妮, 等. 以内容为中心的下一代宽带网络演进 [J]. 信息通信技术, 2011,5(4):38-43.
- [6] YUAN Hao-wei, SONG Tian, CROWLEY P. Scalable NDN forwarding: concepts, issues and principles [C]//Proc of the 1st International Conference on Computer Communications and Networks. 2012: 1-9.
- [7] CHENG Yi, AFANASYEV A, WANG Lan. Adaptive forwariding in named data networking[J]. ACM SIGCOMM Computer Communication Review, 2012,42(3):62-67.
- [8] 叶润生,徐明伟. NDN 邻居缓存路由策略[J]. 计算机科学与探索,2012,6(7):593-601.
- [9] LI Qing, WANG Dan, XU Ming-wei, et al. On the scalability of router forwarding tables: nexthop-selectable FIB aggregation [C]// Proc of IEEE INFOCOM. 2011:321-325.

(下转第268页)

- [5] CASTELLUCCIA C, ElMALKI K, Bellier L. RFC 5380, Hierarchical mobile IPv6 (HMIPv6) mobility management [S]. [S. l.]: IETF, 2008.
- [6] KIM M S, LEE S K, GOLMIE N. Enhanced fast handover for proxy mobile IPv6 in vehicular networks [J]. Wireless Network, 2012, 18(4):401-411.
- [7] OLIVEIRA L M L, De SOUSA A F, RODRIGUES J J P C. et al. Routing and mobility approaches in IPv6 over LoWPAN mesh networks [J]. International Journal of Communication Systems, 2011,224(11):1445-1466.
- [8] SILVA R, SILVA J. An adaptation model for mobile IPv6 support in LoWPANs, draft-silva-6lowpan-mipv6-00[R]. [S. l.]: Internet Engineering Task Force, 2009.
- [9] KIM J, HAW R, CHO E J, et al. A 6LowPAN sensor node mobility scheme based on proxy mobile IPv6 [J]. IEEE Trans on Mobile Computing, 2012,11(12):2060-2072.
- [10] BAG G, RAZA M, KIM K, et al. LoWMob: intra-PAN mobility support schemes for 6LoWPAN[J]. Sensors, 2009,9(7):5844-5877.
- [11] ISLAM M M, HUH E. Sensor proxy mobile IPv6 (SPMIPv6)—a novel scheme for mobility supported IP-WSNs[J]. Sensors, 2011,11 (2):1865-1887.
- [12] WANG Xiao-nan, ZHONG Shan, ZHOU Rong. A mobility support scheme for 6LoWPAN[J]. Computer Communications, 2012,35 (1):392-404.
- [13] DENKO M K, WEI Chen. A multi-gateway-based architecture for integrating Ad hoc networks with the Internet using multiple foreign agents[J]. International Journal of Ad hoc and Ubiquitous Computing, 2008,3(2):99-109.
- [14] FAN Yan-fei, ZHANG Jing-qiu, SHEN Xue-min. Mobility-aware multi-path forwarding scheme for wireless mesh networks [C]//Proc of IEEE Wireless Communications and Networking Conference. Piscataway: IEEE Press, 2008: 2337-2342.
- [15] KIM H, HONG C, SHON T. A lightweight NEMO protocol to support 6LoWPAN[J]. ETRI Journal, 2008,30(5): 685-695.
- [16] DEVARAPALLI V, WAKIKAWA R, PETRESCU A, et al. RFC

- 3963, Network mobility (NEMO) basic support protocol [S]. [S. l.]; IETF, 2005.
- [17] SHIN M K, KIM H J. L3 mobility support in large-scale IP-based sensor networks (6LoWPAN) [C]//Proc of the 11th International Conference on Advanced Communication Technology. Piscataway: IEEE Press, 2009:941-945.
- [18] WANG Xiao-nan, ZHONG Shan. All-IP communication between wireless sensor networks and IPv6 networks based on location information[J]. Computer Standards & Interfaces, 2013,35(1):65-77
- [19] JARA A J, SILVA R M, SILVA J S, *et al.* Mobile IP-based protocol for wireless personal area networks in critical environments [J]. Wireless Personal Communications, 2011, 61(4):711-737.
- [20] SILVA R, SILVA J S, BOAVIDA F. A proposal for proxy-based mobility in WSNs [J]. Computer Communications, 2012, 35 (10):1200-1216.
- [21] HA M K, KIM D Y, KIM S H, et al. Inter-MARIO: a fast and seamless mobility protocol to support Inter-PAN handover in 6LoWPAN [C]//Proc of IEEE GLOBECOM. Piscataway: IEEE Press, 2010:1-6.
- [22] ZHANG Li-qiang, CHENG Qiang, WANG Ying-ge, et al. A novel distributed sensor positioning system using the dual of target tracking [J]. IEEE Trans on Computers, 2008, 57(2):246-260.
- [23] PATWARI N, HERO A O, PERKINS M, et al. Relative location estimation in wireless sensor networks [J]. IEEE Trans on Signal Processing: Special Issue on Signal Processing in Networking, 2003, 51(8):2137-2148.
- [24] NICULESCU D, NATH B. Ad hoc positioning system (APS) using AoA[C]//Proc of IEEE INFOCOM. 2003;1734-1743.
- [25] PATWARI N, ASH J, KYPEROUNTAS S, et al. Locating the nodes: cooperative localization in wireless sensor networks[J]. IEEE Signal Processing Magazine, 2005, 22(4): 54-69.
- [26] CAMP T, BOLENG J, DAVIES V. A survey of mobility models for Ad hoc network research[J]. Wireless Communications and Mobile Computing, 2002, 2(5): 483-502.

(上接第265页)

- [10] BRODE A, MITZENMACHER M. Network applications of Bloom filters: a survey[J]. Internet Mathematics, 2005,4(1):485-509.
- [11] FENG Yi-hsuan, HUANG Nen-fu, CHEN China-hsing. An efficient caching mechnaism for network-based URL filtering by multi-level counting Bloom filters[C]//Proc of IEEE International Conference on Communications. 2011;1-6.
- [12] ZHOU Zhou, SONG Tian, JIA Yun-de. A high-performance URL lookup engine for URL filtering systems [C]//Proc of IEEE International Conference on Communications. 2010;1-5.
- [13] YUAN Hao-wei, WUN B, CROWLEY P. Software-based implementations of updateable data structures for high-speed URL matching [C]//Proc of the 6th ACM/IEEE Symposium on Architectures for Networking and Communications Systems. 2010.
- [14] ZHAO Xin, LIU Yao-qing, WANG Lan, et al. On the aggregatability of router forwarding tables [C]//Proc of the 29th Conference on Information Communications. Piscataway: IEEE Press, 2010:848-856.
- [15] GHODSI A, SHENKER S, KOPONEN T, et al. Information-centric

- networking: seeing the forest for the trees[C]//Proc of the 10th ACM Workshop on Hot Topics in Networks. New York: ACM Press, 2011.
- [16] ATLAS A, ZININ A. RFC 5286, Basic specification for IP fast-re-route; loop-free alternates [S]. 2008.
- [17] ARIANFAR S, NIKANDER P, OTT J. On content-centric router design and implications [C]//Proc of Re-Architecting the Internet Workshop. New York: ACM Press, 2010.
- [18] WANG Yi, HE Ke-qiang, DAI Hui-chen, et al. Scalable name lookup in NDN using effective name component encoding [C]//Proc of the 32nd IEEE International Conference on Distributed Computing Systems. Washington DC: IEEE Computer Society, 2012;688-697.
- [19] CHLEBUS E, BRAZIER J. Nonstationary Poisson modeling of Web browsing session arrivals [J]. Information Processing Letters, 2007, 102(5):187-190.
- [20] GITTER M, CHERITON D R. An architecture for content routing support in the Internet [C]//Proc of the 3rd USENIX Symposium on Internet Technologies and Systems. 2001;37-48.