

# 命名数据网络 NDN 的域间多路径路由机制

胡晓艳<sup>1,2,3</sup>, 龚俭<sup>1,2,3</sup>

(1. 东南大学 计算机科学与工程学院, 江苏 南京 211189; 2. 江苏省网络技术重点实验室, 江苏 南京 211189;  
3. 计算机网络和信息集成教育部重点实验室, 江苏 南京 211189)

**摘要:**为实现命名数据网络(NDN, name data networking)域间内容互访,提出了一种NDN域间多路径路由机制——MIRNDN。该机制使任意自治系统(AS, autonomous system)仅维护自身及客户AS可达内容的路由信息并聚合路由信息以缓解域间路由的可扩展性问题;采用“无谷底”路由策略引导请求非自身和客户AS内容的Interest报文从多路径探索内容,且请求聚合、网络缓存和自适应转发能优化探索;维护多路径路由信息以支持Interest多路径转发。从理论上分析了MIRNDN机制下FIB大小、路由更新的收敛时间和通信开销,在实际因特网AS级别拓扑上的仿真实验表明MIRNDN缓解了域间路由的可扩展性问题,路由更新的收敛时间较短,通信开销适量,并有效地减少了Interest报文的不必要转发。

**关键词:**命名数据网络;域间路由机制;“无谷底”路由策略;多路径路由;路由聚合

**中图分类号:** TP393

**文献标识码:** A

## Multipath inter-domain routing for named data networking

HU Xiao-yan<sup>1,2,3</sup>, GONG Jian<sup>1,2,3</sup>

(1. School of Computer Science and Engineering, Southeast University, Nanjing 211189, China;

2. Jiangsu Provincial Key Laboratory of Computer Network Technology, Nanjing 211189, China;

3. Key Laboratory of Computer Network and Information Integration Ministry of Education, Nanjing 211189, China)

**Abstract:** To facilitate content retrieval among NDN domains, a multipath inter-domain routing for named data networking (MIRNDN) was proposed. Under MIRNDN, an AS (autonomous system) maintained merely the routing state of content reachable from its own network or from its customers' networks and aggregated routing information to mitigate routing scalability issue; "valley free" routing policy was applied to guide Interest packets that solicit content unreachable from the AS's own network nor from its customers' networks to explore matching data packets via multiple paths, and NDN's requests aggregation, in-network caching and adaptive forwarding optimized such exploration; the reachability information of content from multiple paths was collected to support Interest multipath forwarding. The properties of FIB size and the convergence time and communication cost of routing update under MIRNDN were analyzed in theory. The experimental results in the AS level topology of current Internet validates that MIRNDN mitigates the scalability issue of NDN inter-domain routing, has relative short convergence time and moderate communication cost for routing update, and effectively reduces unnecessary Interest forwarding.

**Key words:** named data networking; inter-domain routing; "valley free" routing policy; multipath routing; routing aggregation

## 1 引言

命名数据网络(NDN, name data networking)<sup>[1-3]</sup>

是全新设计的未来互联网体系结构,具有相对成熟的原型系统<sup>[4]</sup>。与IP网络关注数据内容的位置不同,NDN关注数据内容本身,并使用层次结构的内容名

收稿日期: 2014-12-24; 修回日期: 2015-09-12

基金项目: 国家重点基础研究发展(“973”计划)基金资助项目(2009CB320505); 国家自然科学基金资助项目(60973123); 江苏省科技支撑计划基金资助项目(BE2011173); 江苏省未来网络创新研究院未来网络前瞻性研究基金资助项目(BY2013095-5-03)

**Foundation Items:** The National Basic Research Program of China (973 Program)(2009CB320505); The National Natural Science Foundation of China (60973123); The Scientific and Technological Support Project (Industry) of Jiangsu Province (BE2011173); The Prospective Research Project on Future Network of Jiangsu Future Network Innovation Institution (BY2013095-5-03)

字识别和路由报文。NDN 中数据内容的获取由用户发送 Interest 报文向网络拉取数据内容, Interest 报文的内容名字指定请求的数据内容; 具有匹配名字的数据报文响应 Interest 报文的请求, data 报文包含数据内容的名字、数据本身以及数据生成者的签名。通过显式地命名数据内容并对数据内容签名(即 data 报文内置安全性保护), NDN 具有一系列特征如网络缓存、多路径转发、有状态的自适应转发<sup>[5]</sup>、多播数据交付和数据真实性验证等。

NDN 转发表 FIB 的表项包含一个名字前缀和一个或多个下一跳, 用于指导内容名字与该前缀匹配的 Interest 报文的多路径转发, 因此早期 NDN 配置默认 FIB 表项使 Interest 向所有邻居节点转发, 即相当于简单泛洪。但随着 NDN 网络规模的增长, 泛洪将导致网络中流动大量不必要的流量甚至淹没一些主干链路。因此, 为了能在广域网中有效工作, NDN 需要路由机制为 Interest 报文的转发建立(可能是多条)路径, 使路由器将 Interest 报文往潜在数据源转发, 过滤不必要的 Interest 转发流量。NDN 路由器收集直连网络中内容服务器发布内容的存在性信息并传播给邻接路由器, 邻接路由器将这些内容的可达性信息进一步向其邻接路由器传播, 最终各路由器依据所获知识得到网络中内容的可达性, 计算和安装 FIB 表项。

目前, Dai<sup>[6]</sup>、Torres<sup>[7]</sup>和 Hoque<sup>[8]</sup>等探索了如何在 NDN 域内传播内容名字前缀的可达性。与 IP 网络类似, 随着 NDN 网络规模的增长, 网络中的链路和节点将自然地开始扮演不同的角色, 如某些节点和链路用于提供转送服务; 同一组织管辖的路由器将形成一个自治系统(AS, autonomous system)或叫域, 此时将需要域间路由机制实现 AS 之间内容的互访<sup>[9]</sup>。由于 NDN 尚未大规模部署, 目前鲜有 NDN 域间路由机制设计相关的研究工作。相对于 NDN 域内路由机制, NDN 域间路由机制面临着更严峻的可扩展性挑战, 而且需要支持域间路由策略等。NDN 报文的内容名字具有层次结构, 用“/”将不同层次分开, 可聚合。谷歌已检索到约  $10^{12}$  个 URL, 即使使用前缀聚合且聚合到顶级域名前缀也有约  $2 \times 10^8$  个前缀<sup>[10]</sup>。而其他一些因素如对移动的支持会减少路由可聚合的程度, 因而 NDN 路由器可能需要维护远多于  $2 \times 10^8$  个名字前缀。当前因特网的域间路由和报文转发基于“无谷底”路由策略<sup>[11,12]</sup>而非寻找最短路径。考虑到 NDN 并不试图

改变目前互联网的运行和使用模式, Dibenedetto 等<sup>[9]</sup>认为 NDN 的域间路由应继续遵守“无谷底”路由策略, 使 NDN 域间报文的转发遵从经济诱因, 避免转送对 AS 收入无益的流量。

与 IP 域间路由不同, NDN 域间路由传播的是内容名字前缀的可达性而非 IP 地址前缀的可达性。从 IP 地址前缀切换到内容名字前缀并未降低域间路由可扩展性问题的挑战性, 反而可能使它更具挑战性。当前因特网由全球数千 AS 互连而成, 最大的路由表包含约  $4 \times 10^5$  个 BGP 表项, 这些表项涵盖了  $3.9 \times 10^9$  个 IPv4 地址和  $6 \times 10^8$  台主机<sup>[13]</sup>。即通过基于地址前缀的路由聚合, BGP 路由表项数比 IP 地址数缩小了 4 个数量级。但 NDN 域间路由传播的是内容名字前缀的可达性, 可寻址内容的数量预计高数个数量级, 即使内容名字前缀可聚合, NDN 域间路由需维护的名字前缀比 IP 域间路由需维护的地址前缀仍要高数个数量级。另外, BGP 为每个 IP 地址前缀仅维护一条可达 AS 路径。而为了充分利用 NDN 多路径转发的特性以提高网络传输数据的效率, NDN 域间路由应支持多路径路由。对多路径路由的支持, 一方面进一步增加了 NDN 域间路由需要维护的路由信息; 另一方面, 如果路由需要传播和动态更新每个名字前缀的各条可达路径, 如此多的路由状态加上多路径路由信息的频繁更新, 这可能是路由器技术短期内不可及的。庆幸的是, NDN 自适应的转发平面能监测每个 FIB 表项下一跳获取数据的性能进而快速检测链路的失效和恢复, 进而 NDN 域间路由可以忽略网络中短期的波动如短期的链路失效, 也可以降低对快速收敛的需求。IP 网络中路由扮演主导角色, 转发平面严格遵守路由平面的决策; 相比之下, NDN 中路由由仅辅助转发平面, 引导转发平面有效地获取内容并高效地探测新的或恢复的转发路径, 更多的是传播内容的可达性而不是计算最好的路径<sup>[14]</sup>。

鉴于路由技术短期内难以维护网络中所有内容的路由状态, 且“无谷底”路由策略也可以指导 Interest 报文从多路径探索匹配内容, 本文提出了一种 NDN 域间多路径路由机制 MIRNDN, 该机制具有如下特征: 1) AS 节点仅维护本 AS 网络和客户 AS 网络可达内容的路由状态及动态变化, 并通过路由信息聚合, 进一步缓解域间路由的可扩展性问题; 2) 对于请求本网络或客户网络内容的 Interest

报文, AS 节点依据 FIB 表项确定地将它们转发至本网络或客户网络, 而对于请求非本网络和非客户网络内容的 Interest 报文, AS 节点运用“无谷底”路由策略指导这些 Interest 报文探索式地从多路径去寻找可能的数据源, 而请求聚合、网络缓存以及自适应的转发平面能优化这样的探索; 3) AS 节点为多路径可达的名字前缀维护多个下一跳信息, 鉴于 NDN 的无环路由, 仅向供应者 AS 通告内容的可达性信息并不通告具体的多条可达路径以避免多路径的变化导致频繁的路由更新。需要注意的是, 由于网络缓存的内容可能因为缓存节点的存储空间有限而随时被置换, 波动性比较大, 当前 NDN 路由机制暂不传播缓存内容的可达性。

## 2 背景及相关工作

### 2.1 NDN 背景

NDN 转发平面有 3 个组成部分<sup>[1]</sup>。1) 转发表 FIB, 指导 Interest 报文向匹配 data 报文的潜在数据源转发, 每个表项包含一个名字前缀和下一跳转发的接口, 与 IP 的 FIB 表项的区别是它允许有多个而非一个下一跳。FIB 由网络管理员手工配置或由路由机制动态填充。2) PIT (pending Interest table), 记录已转发但匹配 data 报文尚未返回的 Interest 报文以及 Interest 报文进来的接口, 以便 data 报文原路返回请求者。3) CS (content store), 缓存途经 data 报文以响应其他用户后续对该数据的请求, 即网络缓存。

当 Interest 报文到达时, 路由器依据 Interest 的内容名字首先查看 CS 是否已缓存请求的内容。若是, 直接向 Interest 进来的接口返回存储的副本。否则, 查看 PIT 以核实请求该内容的 Interest 是否已向上游节点转发。若已发, 该 Interest 被丢弃, 同时为了避免路由环路(无环路由), 节点仅在核实 Interest 的 nonce 字段值(Interest 发送者生成的随机数以区别于其他请求者)未在匹配 PIT 表项中记录时才记录其 nonce 字段值及其到达的接口。这被称为请求聚合, 因为多个请求相同内容的 Interest 报文中只有第一个到达的会向上游转发。若没有匹配 PIT 表项, 对内容名字以最长匹配方式查找匹配 FIB 表项并将 Interest 向匹配 FIB 表项指定的下一跳接口转发, 可能多个转发接口即支持多路径转发。若无匹配 FIB 表项, Interest 被丢弃。匹配 data 报文依据沿途路由器 PIT 表项记录的 Interest 报文

进入的接口信息原路返回请求者(同时沿途节点中匹配 PIT 表项删除), 且可以被沿途路由器缓存。通过 Interest 和 data 报文的对称路径路由和路由器中 PIT 表项维护的状态信息能观察到 FIB 表项提供的各下一跳数据交付的性能如时延大小、链路失效等以优化后续 Interest 的转发, Yi 等<sup>[5]</sup>已对 NDN 有状态的自适应转发开展了研究。

### 2.2 “无谷底”路由策略

由于经济诱因, 在当前因特网的运行和使用中, AS 节点设置了“无谷底”路由策略, 引导网络流量的流动遵从经济诱因以避免转送对 AS 收入无益的流量。即 AS 为客户 AS 提供转送服务, 允许客户 AS 访问外部因特网以及客户 AS 被外部网络访问; 某些情况下, AS 之间以对等 AS 的形式互联, 相互允许对方访问本网络及客户网络; 但 AS 不会为 2 个供应者 AS 之间、2 个对等 AS 之间或供应者 AS 与对等 AS 之间提供转送服务。具体地, AS 节点可以将客户 AS 收到的报文向除进来的接口外的所有接口包括供应者 AS、客户 AS 和对等 AS 的接口转发, 因为作为供应者时, AS 节点有偿为客户转发流量, 应该为客户提供可靠的网络访问服务; 对于从供应者 AS 或对等 AS 收到的报文, AS 节点仅可以将它向客户 AS 转发, 因为若节点将它向其他供应者 AS 或对等 AS 转发, 这些供应者 AS 或对等 AS 并不会为此流量向该节点付费, 该节点反而需为此向其他供应者 AS 付费或影响与其他对等 AS 间的流量平衡, 该经济抑制因素导致 AS 节点不应向其他供应者 AS 或对等 AS 转发来自于供应者或对等 AS 的报文。

因此, NDN 域间路由可以运用“无谷底”路由策略“过滤”网络中由于经济刺激因素而导致的非必要向供应者 AS 和对等 AS 转发的 Interest 流量。此外, 与 NDN 多路径转发特性和自适应转发平面结合, “无谷底”路由策略可以指导 Interest 报文从多路径探索数据源, 这个数据源可能是内容发布者也可能是网络缓存节点; 即在过滤不必要 Interest 转发流量的同时保持探索内容的能力, 同时利用 NDN 网络缓存的优势。

### 2.3 相关工作

目前, NDN 路由相关的研究比较有限, 其中, Dai<sup>[6]</sup>、Torres<sup>[7]</sup>和 Hoque<sup>[8]</sup>等研究了 NDN 域内路由机制。2 层域内路由协议<sup>[6]</sup>表面上和 NLSR<sup>[8]</sup>相似, 但它们在以下几方面不同。首先, 2 层路由协议用

OSPF 收集拓扑信息和计算最短路径，而 NLSR 用 SYNC 协议而非泛洪的方式传播链路状态广播 LSA。其次，2 层路由协议的路由信息不是以 Interest/data 形式发送，因此它不享受带签名的更新带来的安全性。再次，2 层路由协议的多路径转发仅限于数据内容有多个内容提供者的情况，但 NLSR 还支持将请求多路径转发至同一内容提供者。在基于控制器的路由机制 CRoS<sup>[7]</sup>下，控制器存储网络拓扑，计算路由并存储命名数据的位置信息，进而为网络中任何命名数据安装路由。但 CRoS 中网络泛洪特殊格式的 Interest 报文以寻找控制器可能导致高开销。

NDN 域间路由相关研究有 3 项。其中，文献[15]将 BGP 协议简单扩展为支持内容名字的通告，未考虑 NDN 域间路由面临的实际问题，如路由的可扩展性问题，对多路径路由的支持等。NDN 项目组提出的双曲路由<sup>[16]</sup>和基于 ISP 的名字聚合路由机制<sup>[3]</sup>主要考虑解决 NDN 域间路由的可扩展性问题。双曲路由是一种几何路由机制，依赖双曲坐标系以高效地通过网络发送报文。双曲路由假设存在一种机制可以获取名字前缀和路由器在双曲坐标系中的坐标，Interest 报文途经的路由器首先计算邻居路由器和 Interest 报文名字前缀的坐标距离，然后贪心地向坐标距离中最小的邻居路由器转发 Interest 报文。双曲路由具有高可扩展性，因为节点无需维护路由表或 FIB，而且除了学习邻居节点的坐标外，没有动态的路由更新。但是，双曲路由在 NDN 中的可行性取决于基础拓扑是否是双曲线的，这是个尚未探索的问题；另外，作为域间路由，双曲路由如何支持路由策略是另一个等待探索的问题；此外，双曲路由中 Interest 报文的转发可能陷入局部最优的情况，限制了它的使用。基于 ISP 的名字聚合路由机制中，路由器仅需维护 ISP 级别具有层次结构的名字前缀，用户的数据访问首先需要额外的映射基础设施将用户选择的内容名字映射为 ISP 分配的名字，然后，网络中的路由器依据报文中 ISP 分配的名字路由报文。因此，该路由机制需要搭建额外的映射基础设施，而且映射的查询给数据的访问带来了额外的开销。由于 NDN 项目组对于双曲路由和基于 ISP 的名字聚合机制在 NDN 中的可行性及具体使用细节尚不明确，将内容名字解析为坐标或映射为 ISP 分配的名字均会产生额外的访问时延，加上双曲路由难以支持域间路由策略。本文提

出了支持路由策略的 NDN 域间多路径路由机制 MIRNDN，并从理论和实验 2 方面验证其可行性。MIRNDN 机制支持域间路由策略，通过仅维护网络中部分内容的路由状态及聚合路由信息的方式缓解域间路由的可扩展性问题，并维护多路径路由信息以支持 Interest 报文的的多路径转发。

### 3 多路径路由机制

NDN 域间路由的最终目的是为各 AS 节点安装正确的 FIB。由于 FIB 指导 Interest 报文向潜在数据源转发，过滤不必要的 Interest 报文转发，本文将 FIB 形象地称为“过滤器”。MIRNDN 生成 2 级“过滤器”：第一级“过滤器”依据 AS 节点收集的本网络提供内容的可达性信息（由域内路由机制如 NLSR 获取，不在本文讨论范围内）以及客户 AS 节点通告的可提供内容的可达性信息一起建立 FIB，每个 FIB 表项可能包含多个下一跳，指导请求本网络或客户网络内容的 Interest 报文的的多路径转发；第二级“过滤器”依据“无谷底”路由策略建立，指导在第一级“过滤器”中未找到匹配 FIB 表项的 Interest 报文（即 Interest 请求内容在本网络和客户网络均无法获取）探索式地从多路径去寻找内容，而请求聚合、网络缓存和自适应转发会优化这样的探索。这 2 级“过滤器”的具体格式如图 1 所示。

名字前缀	下一跳接口列表
$p_1$	下一跳接口列表 1
$p_2$	下一跳接口列表 2
$p_3$	下一跳接口列表 3
...	...

(a) 依据路由更新生成的 FIB

到达接口列表	下一跳接口列表
客户 AS 接口列表	供应者和对等 AS 接口列表

(b) 依据“无谷底”路由策略生成的 FIB

图 1 2 种 FIB 格式

MIRNDN 下 AS 级别网络拓扑中的链路被分类为客户 AS/供应者 AS 链路或对等 AS 链路。与当前因特网运行类似，假设没有任何中央控制，链路(A, B)的分类是 AS 节点 A 和 B 间的局部决策，无需与任何中心位置协调或报告。据此，AS 节点的每个接口被标识为客户节点、供应者节点或对等节点，该信息由节点操作员手工配置。AS 节点间关系在实际情

况下应相对稳定，接下来的描述均假设 AS 间的关系固定且基于下面的符号：AS 节点  $X$  的供应者 AS 集合是  $Pro(X)$ ，对等 AS 集合是  $Peers(X)$ ，客户 AS 集合是  $Cus(X)$ 。另外，名字前缀间的关系定义如下。

**定义 1 (覆盖关系):** 较短的父前缀(covering)  $p1$  完全包含较长的子前缀(covered)  $p2$  的命名空间。若  $p1$  的层次数为  $n$ ，则  $p2$  的前  $n$  个层次与  $p1$  一致。用符号表示  $p1$ 、 $p2$  的覆盖关系为  $p2 \subset p1$ ，则有  $\frac{p2}{n} = p1$ 。

**定义 2 (公共前缀):** 层次数分别为  $m$  和  $n$  的 2 个名字前缀  $p1$ 、 $p2$ ，若存在最大的整数  $k$ ， $k \leq m$  且  $k \leq n$ ，使  $\frac{p1}{k} = \frac{p2}{k}$ ，那么  $p1$  和  $p2$  有层次数为  $k$  的公共前缀。

### 3.1 基于路由更新的“过滤器”

MIRNDN 中客户 AS 向供应者 AS 发送路由更新信息  $update$ ，通知供应者 AS 通过它可达或不再可达的内容的名字前缀，供应者 AS 据此建立 FIB 表项以过滤向客户 AS 转发的 Interest 流量。运行 MIRNDN 的 AS 节点的路由器维护 3 种路由数据信息：adj-ribs-in 存储各客户 AS 通告的名字前缀；FIB 是根据 adj-ribs-in 生成并安装在本地的转发表；根据本地输出策略，AS 节点有选择地向各供应者 AS 通告可达内容的前缀，而 adj-ribs-out 存储 AS 节点向各供应者通告的前缀。图 2 给出了 AS  $X$  的路由器在接收到路由更新  $update$  后的处理步骤。首先更新 adj-ribs-in，然后重新计算和安装 FIB，其中，可能涉及 FIB 表项的聚合工作，接着根据本地输出策略更新 adj-ribs-out，最后选择性地各供应者 AS 发送路由更新信息。

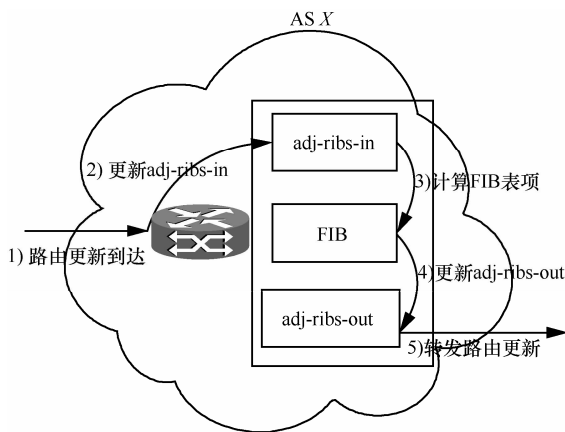


图 2 AS  $X$  的路由器对路由更新的处理过程

基于客户 AS 路由更新生成的 FIB 的具体格式如图 1(a)所示，FIB 表项记录具有某个名字前缀的 Interest 流量可以转发的一个或多个下一跳接口。网络中内容的可达性状态是动态变化的，这些变化需以路由更新的形式向供应者 AS 传播，通知供应者当前可达或不再可达的内容的名字前缀，即路由更新可能包含路由通告 (advertisement) 和路由撤销 (withdrawal)，具体格式如图 3 所示。路由通告通知新可达内容的名字前缀，由数据源所在 AS 节点生成。路由撤销已通告的名字前缀，表明具有该名字前缀的内容不再可达，可能是内容提供者所在 AS 不再提供以该名字前缀命名的内容而主动发送，也可能是中间节点检测到相邻客户 AS 的故障而被动撤销该故障节点所通告的名字前缀。若为后者，链路故障达到一定时间后，路由撤销才生成，因为自适应转发平面允许路由忽略网络中短期的波动。本文将最初生成路由更新信息的 AS 称为 origin。路由撤销可以立即转发，但为了避免过多的路由更新，相邻路由通告的发送之间具有最小时间间隔  $minrouteadver$ 。路由更新在 AS 层次的拓扑中仅从下往上传播，因而路由更新的传播不应出现环路。

路由通告长度
$p1$
$p2$
...
路由撤销长度
$p1'$
$p2'$
...

图 3 路由更新报文的内容格式

图 4 和图 5 分别给出了 AS 节点处理收到的路由通告和路由撤销的实例，两图中实线箭头表示供应者和客户 AS 间的关系，虚线箭头表示路由更新发送的路线。图 4 中 AS  $W$  向供应者 AS  $Y$  发送了 2 个名字前缀  $com/google/map$  和  $com/google/sports$  的路由通告。当收到该路由通告时， $Y$  在 adj-ribs-in 中  $W$  通告的前缀集合中添加这 2 个前缀；由于  $com/google/map$  和  $com/google/sport$  拥有的公共前缀  $com/google$  且均来自客户  $W$ ， $Y$  为公共前缀  $com/$

google 生成一个 FIB 表项; 然后  $Y$  根据本地输出策略决定将  $com/google$  通告给供应者  $X$ , 在  $minrouteadver$  超时后在  $adj\_ribs\_out$  中注明向供应者  $X$  通告前缀  $com/google$  并生成路由通告发送给  $X$ 。内容可能存在多个数据源, 同一数据源也可能有多条路径可达, 如图 4 中  $X$  收到另一客户 AS  $Z$  关于前缀  $com/google$  的路由通告。基于此,  $X$  将这 2 个路由通告的信息聚合起来生成一个 FIB 表项, 指出与该前缀匹配的 Interest 流量可以向 2 个客户 AS  $Y$  和  $Z$  转发。图 5 中 AS  $Z$  向供应者 AS  $Y$  发送关于前缀  $com/google$  的路由撤销。收到该路由撤销后,  $Y$  从  $adj\_ribs\_in$  中关于  $Z$  通告的前缀集合中删除  $com/google$ , 并将  $Z$  从  $com/google$  的 FIB 表项的下一跳集合中删除。由于仍可以从另一客户 AS  $W$  处访问具有名字前缀  $com/google$  的内容,  $Y$  不用向供应者 AS  $X$  发送  $com/google$  的路由撤销信息。

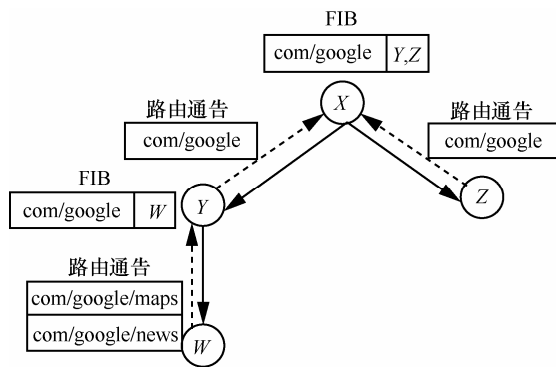


图 4 路由通告处理实例

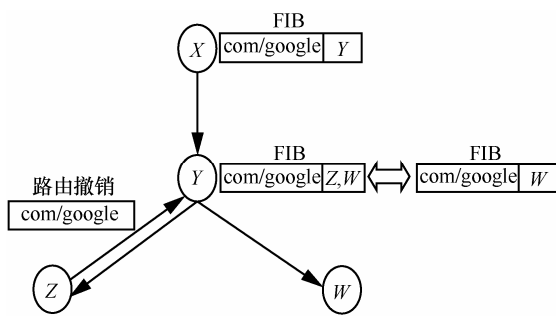


图 5 路由撤销处理实例

过程 1 描述 AS  $X$  收到客户 AS  $Y$  发送的关于名字前缀  $P$  的路由更新  $update$  时启动的路由更新处理过程。其中, 第 1) 步更新  $adj\_ribs\_in$ , 在  $adj\_ribs\_in$  中关于  $Y$  通告的前缀集合中添加 ( $update$  是路由通告时) 或删除 ( $update$  是路由撤销时)  $P$ 。第 2) ~ 25) 步表示若  $update$  是路由通告时  $X$  的更新过程。第 3) 和第 4) 步中,  $X$  查看 FIB 中是否存在关于  $P$  的表

项  $entry$ 。若存在, 说明有其他客户 AS 通告了该前缀且该前缀已通告  $Pro(X)$  中供应者, 无需再向供应者通告, 仅需在相应 FIB 表项的下一跳集合中添加  $Y$ 。否则, 第 5) ~ 8) 步中,  $X$  查看若存在覆盖  $P$  的 FIB 表项, 那么  $X$  无需向  $Pro(X)$  中供应者通告  $P$ ; 但是若覆盖  $P$  的 FIB 表项的下一跳集合不包含  $Y$ ,  $X$  仍需为  $P$  建立一个 FIB 表项。若不存在覆盖  $P$  的 FIB 表项, 第 10) ~ 12) 步中,  $X$  查看若存在 FIB 表项和  $P$  拥有公共前缀且均来源于  $Y$  发送的路由通告,  $X$  用公共前缀聚合生成新 FIB 表项。否则第 14) 步中,  $X$  直接为  $P$  生成一个新的 FIB 表项, 这种情况下  $X$  中关于  $Y$  通告的被  $P$  覆盖的 FIB 表项可以删除 (第 15) ~ 17) 步)。对于后 2 种情况,  $X$  具有新名字前缀需向供应者 AS 通告。第 19) ~ 24) 步中,  $X$  等待  $minrouteadver$  超时, 生成新路由通告, 根据本地输出策略向供应者发送新路由通告并更新相应的  $adj\_ribs\_out$ 。过程 1 中第 26) ~ 37) 步表示若  $update$  是路由撤销时  $X$  的更新过程: 第 27) ~ 28) 步中, 若  $X$  存在关于  $P$  的 FIB 表项  $entry$ , 或者存在覆盖  $P$  的 FIB 表项  $entry$  且  $Y$  没有通告其他被该  $entry$  覆盖的名字前缀, 则从  $entry$  的  $next-hop$  中删除  $Y$ ; 第 29) ~ 36) 步中, 若上述因为该路由撤销而被更新的 FIB 表项  $entry$  的  $next-hop$  为空集, 该  $entry$  前缀所标识内容的可达性需向供应者撤销,  $X$  立即生成新的路由撤销信息, 根据本地输出策略发送给供应者并更新相应  $adj\_ribs\_out$ 。

过程 1 MIRNDN 的路由更新处理过程

/\*  $X$  对客户  $Y$  发送的关于名字前缀  $P$  的路由更新  $update$  的处理过程 \*/

- 1) Update\_adj\_ribs\_in (update, Y) // 更新 adj\_ribs\_in
- 2) IF Is\_advertisement (update) // update 是路由通告信息
- 3) IF entry=Lookup\_entry(P) // 存在关于  $P$  的 FIB 表项 entry, 在 entry 的 next-hop 中添加  $Y$
- 4) entry.next-hop+= {Y}
- 5) ELSEIF entry=Lookup\_Covering\_entry(P) //  $P \subset entry.ContentPrefix$
- 6) IF  $Y \notin Yentry.next-hop$  // 覆盖  $P$  的 FIB 表项 entry 的 next-hop 不包含  $Y$
- 7) entry' = Create\_entry(P, Y)
- 8) ENDIF
- 9) ELSE // 为  $P$  创建一个新的 FIB 表项 entry'

并向供应者 AS 转发路由通告

10) IF  $entry=Lookup\_CommonPrefix\_entry(P)$  and  $entry.next-hop==\{Y\}$

//  $\frac{entry.ContentPrefix}{k} = \frac{p}{k}$  且它们具有相同的

下一跳  $Y$

11)  $entry' = Aggregate\_entry(P, entry.ContentPrefix, Y)$  //聚合成一个新的 FIB 表项

12) Delete\_entry( $entry$ )

13) ELSE

14)  $entry' = Create\_entry(P, Y)$

15) IF  $entry=Lookup\_Covered\_entry(P)$  and  $entry.next-hop==\{Y\}$ // $entry.ContentPrefix \subset P$  且下一跳也为  $Y$

16) Delete\_entry( $entry$ )

17) ENDIF

18) ENDIF

19) wait\_minrouteadver // 等待 minrouteadver 超时

20)  $Adver=Create\_advertisement(entry'.ContentPrefix)$  // 创建新路由通告

21) FOREACH  $Z \in Pro(X)$  and  $X.export\_policy(Z, adver, Y) == true$

//依据本地输出策略向供应者发送  $adver$  并更新  $adj\_ribs\_out$

22) update\_adj\_ribs\_out( $adver, Z$ )

23) Send\_update( $Z, adver$ )

24) ENDFOR

25) ENDIF

26) ELSE //update 是路由撤销信息

27) IF  $entry=Lookup\_entry(P)$  or ( $entry=Lookup\_Covering\_entry(P)$  and  $Has\_adj\_ribs\_in(entry.ContentPrefix, Y) == false$ )

//存在关于  $P$  的 FIB 表项  $entry$ , 或者存在覆盖  $P$  的 FIB 表项  $entry$  且  $Y$  未通告其他被  $entry$  覆盖的前缀

28)  $entry.next-hop=\{Y\}$

29) IF  $entry.next-hop==\emptyset$  // $entry$  的  $next-hop$  为空集, 删除  $entry$

30) Delete\_entry( $entry$ )

31)  $Withdraw=Create\_withdrawal(entry.ContentPrefix)$  //创建新的路由撤销

32) FOREACH  $Z \in Pro(X)$  and

$X.export\_policy(Z, Withdraw, Y)==true$

//依据本地输出策略向供应者发送  $Withdraw$  并更新  $adj\_ribs\_out$

33) Update\_Adj\_Ribs\_Out( $Withdraw, Z$ )

34) Send\_Update( $Z, Withdraw$ )

35) ENDFOR

36) ENDIF

37) ENDIF

38) ENDIF

### 3.2 基于“无谷底”路由策略的“过滤器”

若 AS  $X$  收到的 Interest 报文请求的内容不在本网络和客户 AS 网络 (即在第一级“过滤器”中未找到匹配 FIB 表项),  $X$  依据“无谷底”路由策略建立的 FIB 表项指导 Interest 报文的转发。MIRNDN 下 AS 节点根据各接口的角色遵从“无谷底”路由策略建立一条 FIB 表项如图 1(b)所示。

依据“无谷底”路由策略, AS  $X$  可以将将从客户收到的 Interest 报文向除进来的接口外的所有接口包括  $Pro(X)$ 、 $Peers(X)$  和  $Cus(X)$  中 AS 的接口转发。但从第一级“过滤器”核实的结果已表明客户 AS 网络均没有请求的内容, 故如图 1(b)中 FIB 表项所示, 仅将供应者 AS 和对等 AS 的接口作为这些来自客户 AS 的 Interest 报文去探索潜在数据源的选项。

另外, 依据“无谷底”路由策略, 对于从供应者 AS 和对等 AS 收到的 Interest 报文, 节点  $X$  仅可以将它向客户 AS 转发。与此同时, 从第一级“过滤器”的核实结果已表明客户 AS 网络均没有要请求的内容。综合上述 2 点, 这些来自供应者 AS<sup>[1]</sup> 和对等 AS 的 Interest 报文应该被丢弃, 也即没有匹配的 FIB 表项。

### 3.3 Interest 报文的转发过程

MIRNDN 下 Interest 报文转发是上述 2 级“过滤器”综合作用的结果, 如过程 2 所述 (关于  $X$  从邻接节点  $Y$  收到内容名字包含  $prefix$  的 Interest 报文  $int$  且  $int$  在  $X$  的 CS 和 PIT 均无匹配表项时  $int$  的转发。请注意 CS 网络缓存和 PIT 请求聚合的特征可以优化内容的探索)。若  $int$  在节点  $X$  的第一级“过滤器”中经过最长匹配后找到匹配 FIB 表项  $entry$  (过

注1 一般来说, Interest 来自于供应者且在第一级“过滤器”未找到匹配表项的情况应不存在。在第一级“过滤器”未找到匹配表项说明本节点未向供应者通告或撤销了相应前缀下内容的可达性, 供应者不应将这样的 Interest 转发到本节点, 但不排除路由未收敛时出现此情况。

程 2 的第 1)~4) 步), 说明从客户 AS 能够获取用户请求的内容, 则依据具体转发策略 (自适应转发), 将 *int* 向 *entry* 的 *next-hop* 中所有客户 AS 接口转发, 或向 *next-hop* 中部分接口转发, 进而客户 AS 仅接收它们路由更新中声称可达内容的 Interest 流量。若 *int* 在第一级 “过滤器” FIB 中未找到匹配表项, 其转发则依据第二级 “过滤器” 即 “无谷底” 路由策略生成的 FIB 表项, 分 2 种情况: 若 *Y* 是客户 AS, 依据具体转发策略, *X* 可以将 *int* 转发到所有对等 AS 和供应者 AS 或它们的子集(过程 2 的第 6)~10)步); 若 *int* 来自于对等 AS 或供应者 AS, 节点丢弃 *int*。

过程 2 MIRNDN 下 Interest 报文的转发

/\* *X* 从 *Y* 收到 *int*, 其内容名字是 *prefix* \*/

- 1) IF *entry*=LPM\_entry(*prefix*) //在第一级 “过滤器” 最长匹配后找到与 *prefix* 匹配的 FIB 表项 *entry*
- 2) FOREACH *Z* ∈ *entry.next-hop* and Forward\_strategy(*entry.ContentPrefix*, *Z*)==true
- 3) Forward(*Z*, *int*) //向 *Z* 转发 *int*
- 4) ENDFOR
- 5) ELSE //依据第二级 “过滤器” 转发 *int*
- 6) IF *Y* ∈ *Cus*(*X*)//*Y* 是客户 AS, 向供应者 AS 和对等 AS 转发 *int*
- 7) FOREACH *Z* ∈ *Peers*(*X*) ∪ *Pro*(*X*) and Forward\_strategy(*entry.ContentPrefix*, *Z*)==true
- 8) Forward(*Z*, *int*) //向 *Z* 转发 *int*
- 9) ENDFOR
- 10) ENDIF
- 11) ENDIF

图 6 举例展示了 MIRNDN 下一个 AS 级别拓扑图上 Interest 报文的转发路径 (如图 6(a)所示) 和匹配 data 的返回路径 (如图 6(b)所示)。其中, 向下的实线箭头表示供应者 AS 和客户 AS 间的关系边, 直的实线表示对等 AS 间的关系边。图中 AS C 的用户发布了以某名字前缀命名的内容, 该内容尚未在网络中缓存, C 将此内容的可达性通告给供应者 E (用向上的实线箭头 C 至 E 表示), 其中, E 可将此路由通告发送给它的供应者(图中 E 没有供应者), 但并不将此通告发送给它的对等节点(D 或 F)以及客户节点(B)。AS A 的某用户试图访问 AS C 的内容, 2 个子图中的虚线箭头分别组成了 Interest 报文和 data 报文途经的路径。

如图 6(a)所示, A 将 Interest 报文发送给供应者 D。因为 D 的客户 B 发送的路由更新信息表明 B 没

有请求的内容, 所以 D 不会将 Interest 转发给 B; 但因为该 Interest 报文来自于客户 A, 依据 “无谷底” 路由策略, D 可将其转发给供应者 AS (D 没有供应者) 和对等节点 (E 和 F)。请注意虽然 D 发送了一个不必要的 Interest 报文给对等 AS F (因为 F 没有请求的内容), 但 “无谷底” 路由策略避免 F 进一步将该 Interest 转发给 F 另一对等节点 E (同理 E 不向 F 转发该 Interest)。而 E 会将 Interest 转发给客户 C, 因为 C 发送给 E 的路由更新包含了所请求内容的可达性 (但 E 不会向客户 B 转发该 Interest, 因为 B 的路由更新未声明它能提供请求的内容)。最后, 如图 6(b)所示, 匹配的 data 报文从节点 C 沿着 Interest 报文沿途建立的 PIT 表项信息原路返回到请求者 A。

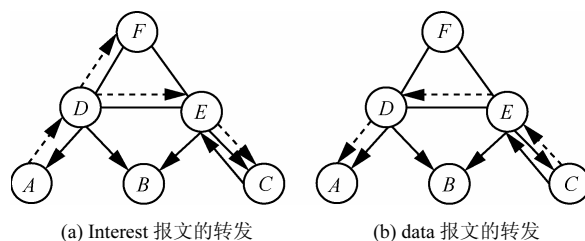


图 6 MIRNDN 下 Interest 和 data 报文的转发

从图 6 的例子可以看到, MIRNDN 的 2 级 “过滤器” 有效地指导 Interest 报文的转发, 同时避免不必要的 Interest 报文转发 (从 D 到 B, F 到 E 或 E 到 F, 以及 E 到 B)。同时, 由于 Interest 报文沿着 2 条路径寻找请求的内容 (多路径路由), 如果 F 缓存了请求的内容, 相对于 C 更近的 F 则会用缓存的内容响应该请求, 进而提高内容获取的性能。

## 4 分析讨论

### 4.1 AS 节点 FIB 表项数目

MIRNDN 下路由更新仅从客户 AS 向上层供应者 AS 发送, 没有任何路由信息向下层客户 AS 发送, 对等节点之间也没有路由信息的交换。AS 节点的 FIB 维护以自身为根和客户网络一起组成的子树中可达内容的路由信息。AS 节点的 FIB 表项集合是本节点和客户节点所提供内容路由信息的并集, 即  $FIB(X) = \{\cup_{c \in Cus(X)} Route(c) \cup Route(X)\}$ 。

其中,  $FIB(X)$  表示节点 X 的 FIB 表项集合,  $Route(X)$  表示 X 本身提供内容的路由信息,  $Cus(X)$  表示 X 所有客户 AS 集合,  $Route(c)$  表示 X 的客户之一 c 向 X 通告的可达内容的路由信息。在该模型下, 根

网络的 FIB 仅列举它自身所提供内容的路由信息; tier 1 网络节点的 FIB 则列举自身及所有客户 AS 可达内容的路由信息, 这个 FIB 可能很大。但 tier 1 AS 的客户节点需为它在 tier 1 网络节点中的 FIB 表项付费, 因此 FIB 大小与供应者 AS 的客户群大小直接相关, 也即收益与此相关。

#### 4.2 路由更新的收敛时间和通信开销

MIRNDN 的路由收敛是指路由通告中新名字前缀下内容的可达性最终传播至路由通告的 origin 的任意上层供应者 AS 或路由撤销所撤销名字前缀下内容不可达的事实被 origin 的所有上层供应者 AS 接受。路由收敛经历的时间和需发送的更新报文数分别被称为路由更新的收敛时间和通信开销。收敛时间与路由更新周期  $\text{minrouteadver}$  有关。假设路由更新信息在相邻 2 个 AS 之间传输的平均延迟为  $h$ , 在每个 AS 节点被延迟的平均时间为  $\delta$  (包括在节点的排队延迟和处理时间),  $h, \delta \ll \text{minrouteadver}$ , AS 级别拓扑的直径为  $d$ 。路由更新  $\text{update}$  (关于名字前缀  $\text{prefix}$ ) 的 origin 是  $S$ ,  $S$  可能拥有多个供应者 AS。 $\text{update}$  可能沿着多条路径向  $S$  的 tier 1 AS 传播, 这些路径上的 AS 组成了  $S$  的上层供应者 AS 集合  $V_S$ 。 $V_S$  与集合中节点间客户、供应者的关系边集合  $E_S$  形成了 AS 级别拓扑的一个子图  $G_S(V_S, E_S)$ , 被称为  $S$  的上层关系图,  $d_S$  是该子图的直径。

**定理 1** 路由通告  $\text{update}$  的收敛时间上限是  $O((\text{minrouteadver}+h+\delta)d_S)$ 。

**证明** 路由通告  $\text{update}$  向 origin 的所有上层供应者传播直到 origin 的所有 tier 1 供应者。当上层供应者  $P \in V_S$  首次收到  $\text{update}$  时(沿最短路径到达),  $P$  则得知客户可以提供新内容, 此时  $\text{update}$  在  $P$  处已收敛(若  $\text{update}$  后续从其他路径到达  $P$ , 这只是为内容的获取提供新路径)。因此路由通告  $\text{update}$  的收敛时间是  $\text{update}$  从 origin  $S$  沿着最短路径到达  $V_S$  中任一 tier 1 AS 所需时间中的最长者。由于最短路径中的最长者上限为  $d_S$  且路由通告受  $\text{minrouteadver}$  限制, 路由通告  $\text{update}$  收敛时间上限为  $O((\text{minrouteadver}+h+\delta)d_S)$ 。□

**定理 2** 路由通告  $\text{update}$  收敛期间发送的更新报文数上限是  $O(|E_S|)$ ,  $|E_S|$  是  $E_S$  集合大小。

**证明** 对于路由通告信息  $\text{update}$ , 上层供应者  $P \in V_S$  仅在首次收到  $\text{update}$  且  $P$  之前没有向供应者发送关于或覆盖  $\text{prefix}$  的路由通告时,  $P$  才会向通

往供应者 AS 的链路发送  $\text{update}$ 。因此路由通告  $\text{update}$  在  $G_S$  的每条边上最多引发一次路由更新的发送, 即路由通告  $\text{update}$  收敛期间发送的更新报文数上限是图  $G_S$  的边数  $O(|E_S|)$ 。

**定理 3** 路由撤销  $\text{update}$  的收敛时间上限为  $O((h+\delta)|V_S|)$ ,  $|V_S|$  是  $V_S$  集合的大小。

**证明** 路由撤销  $\text{update}$  可能沿多条路径到达上层供应者  $P \in V_S$ , 仅当该撤销信息从这多条路径均到达  $P$  时,  $\text{prefix}$  下内容的可达性才可能会被  $P$  撤销并进一步向  $P$  的供应者 AS 通报。为了让路由撤销  $\text{update}$  中  $\text{prefix}$  下内容不可达的事实被  $S$  的所有上层供应者 AS 接受, 它的收敛时间上限为  $\text{update}$  从  $S$  经最长路径到  $S$  任一 tier 1 AS 所需时间中的最长者。由于路由撤销  $\text{update}$  的传播路径沿着 AS 的层次从下向上, 不应出现环路, 最长路径的上限是  $|V_S|$ 。因此路由撤销  $\text{update}$  的收敛时间上限为  $O((h+\delta)|V_S|)$ 。

**定理 4** 路由撤销  $\text{update}$  收敛期间发送的更新报文数上限是  $O(|E_S|)$ 。

**证明** 路由撤销  $\text{update}$  收敛期间, 仅当  $\text{update}$  从所有路径均到达  $P \in V_S$  时,  $P$  才有可能唯一也是最多一次将此撤销信息向通往供应者 AS 的链路发送, 因此路由撤销  $\text{update}$  在  $G_S$  的每条边上最多引发一次路由撤销的发送, 故路由撤销  $\text{update}$  收敛期间发送的更新报文数上限是  $O(|E_S|)$ 。

从上述分析看到, MIRNDN 下路由更新在 origin  $S$  的上层关系图的每条链路上最多引发一次路由更新报文的发送, 路由收敛所需时间也较短。此外, 即使路由撤销信息尚未收敛, 自适应的转发平面可能已发现撤销的数据源不可达, 且多路径路由为内容的获取提供了额外的健硕性。

#### 4.3 安全性

MIRNDN 下节点间交换的路由更新报文携带的有效签名信息基于 NDN 的信任机制, 路由更新信息的真实性可以进行验证, 因而路由信息可靠安全, 可以避免前缀劫持这样的攻击。但 NDN 存在 Interest 泛洪攻击, 即攻击者可能发送一些 Interest 报文, 请求大量实际上在网络中不存在的内容, 这样的 Interest 的多路径探索没有匹配的内容返回, 但它耗费了网络带宽以及网络节点为其 PIT 分配的有限资源。而在 MIRNDN 下, 若 Interest 报文请求的内容不属于本网络和客户网络, AS 节点按照“无谷底”路由策略从多路径去探索内容, 这时

可能出现 Interest 泛洪攻击。MIRNDN 并未引入新攻击，而且 NDN 自适应转发平面可以及时识别并过滤 Interest 泛洪攻击流量，将攻击的危害降低至最低<sup>[17]</sup>。

## 5 性能评估

本文用 C 语言编写了一个简易版本的 MIRNDN，在一个带注释的 AS 级别因特网拓扑<sup>[18]</sup>中仿真上述路由机制，从 FIB 表大小、路由更新收敛时间和通信开销以及数据平面开销 3 个方面对 MIRNDN 进行评估。实验所用拓扑由 CAIDA 提供，包含 33 508 个 AS 节点和 75 001 条有向边(即链路)；有向边依据客户、兄弟(实验中视为客户/供应者的关系)、对等者以及供应者的关系进行标识<sup>[11,19]</sup>。

本文 5.1 节评估该拓扑中 MIRNDN 下 AS 节点需维护的 FIB 表大小。实验假设每个 AS 网络的内容使用一个公共名字前缀并且 AS 节点向供应者通告本网络内容的名字前缀以及客户 AS 通告的名字前缀<sup>[2]</sup>(这些名字前缀不可聚合，NDN 名字前缀的可聚合性不在本文讨论范围内)，然后计算各 AS FIB 需维护的名字前缀数目，作为评估 MIRNDN 下 AS FIB 表大小的基准。

5.2 节评估 MIRNDN 下该 AS 拓扑中各节点发送的路由通告和路由撤销收敛所需时间及通信开销，具体参数设置如表 1 所示。当一个 AS 节点的路由更新事件发生时，其上层关系图中各供应者节点在当前路由更新周期的时刻点是 0 到  $\text{minrouteadver}$  之间的随机值。各节点路由通告或路由撤销事件的仿真重复 10 遍，然后结果取均值。根据定理 2 和定理 4，本实验以 AS 节点上层关系图中链路的数量作为节点路由更新收敛期间的通信开销上限。

表 1 实验参数设置

参数	设定值
$\text{minrouteadver}$	30 s
$\text{minrouteadver}$ 抖动	$[0.75, 1] \text{ minrouteadver}$
$\Delta$	$[0.01, 0.1] \text{ ms}$
$H$	$[5, 10] \text{ ms}$

注2 用户如何获取内容名字以及 AS 节点命名空间的分配不在本论文的讨论范围之内，本文仅假设存在某种机制(如搜索引擎)使用户可以查到所请求内容的名字以及每个 AS 被分配一个名字前缀。

5.3 节评估 MIRNDN 下各 AS 节点发送的 Interest 报文被转发的次数，以此评估数据平面的开销。即一个 AS 节点发送 Interest 报文请求另一个 AS 提供的内容，各节点按照 FIB 转发该报文，计算 Interest 总共在多少条链路转发。MIRNDN 中 AS 节点主要应用“无谷底”路由策略和客户 AS 的路由更新信息对 Interest 报文的不必要转发进行 2 级过滤。为了分析 MIRNDN 的 2 级“过滤器”各自所发挥的作用，仿真实验首先评估仅使用“无谷底”路由策略，称为没有路由更新的“过滤器”(PPF, pure policy-based filtering)，过滤不必要 Interest 转发的有效性，然后再仿真 MIRNDN 以评估客户 AS 的路由通告在 PPF 基础之上进一步过滤不必要 Interest 转发的有效性。请注意仿真实验中用户所请求内容仅存在于数据源中，尚未在网络中缓存，网络节点也未采用自适应转发而完全按照 FIB 转发 Interest 报文。

### 5.1 FIB 大小

图 7 显示了各 AS 节点 FIB 表项数目的累积分布函数 CDF，其中，近 85% AS(根网络)只需维护一个 FIB 表项，约 95% AS 只需维护小于 10 个 FIB 表项(不到总表项数的 0.03%)，最大的 FIB 有 28 796 个表项，约占总表项数的 89.3%。该拓扑中有 31 个 tier 1 AS 节点，FIB 大小排名前 10 的节点均为 tier 1 节点，接下来的是这些 tier 1 节点的下一级客户 tier 2 节点以及其他 tier 1 节点。这是因为 MIRNDN 中 AS 节点的 FIB 维护由其自身网络及客户 AS 网络组成的子树所提供内容的路由状态，tier 1 AS 节点的网络一般比较大，有的大到甚至它的 tier 2 客户节点也比另一些 tier 1 节点大。值得说明的是，上述关于 FIB 表项数目的实验数据是基于这些名字前缀不可聚合的假设得出的，实际网络环境下节点所通告的名字前缀的可聚合性取决于 NDN 的命名机制和命名管理，应尽可能使供应者提供内容所采用的名字前缀能够覆盖客户网络所提供内容的名字前缀进而可以聚合，因而实际情况下有些节点尤其是网络核心节点的 FIB 大小相对于实验的结果应该小不少。上述实验为评估 MIRNDN 下 FIB 的可扩展性给出了一个参考点，可以看到 FIB 大小是 AS 规模的函数，大部分 AS 所需维护 FIB 较小，只有大的网络核心才需维护较大的 FIB，当每个 AS 有更多内容名字前缀需通告的情况可以此类推，应仍然满足类似关系。

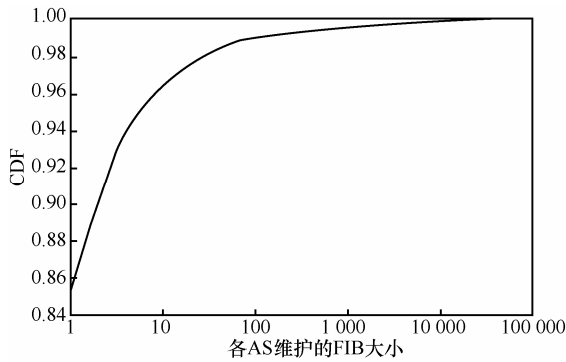


图 7 MIRNDN 下各 AS FIB 大小的 CDF

### 5.2 路由更新收敛时间和通信开销

图 8 给出了各 AS 的路由通告收敛所需时间的 CDF。可以看到，在实验所用的 AS 拓扑中，MIRNDN 下 AS 节点路由通告的收敛时间在 12~120 s，约 27%节点路由通告的收敛时间小于 40 s，约 70%节点路由通告的收敛时间在 40~80 s，不到 1%节点路由通告的收敛时间大于 100 s。该实验结果说明 MIRNDN 下新数据内容的可达性在发送路由通告后 100 s 之内应该可以被用户访问。

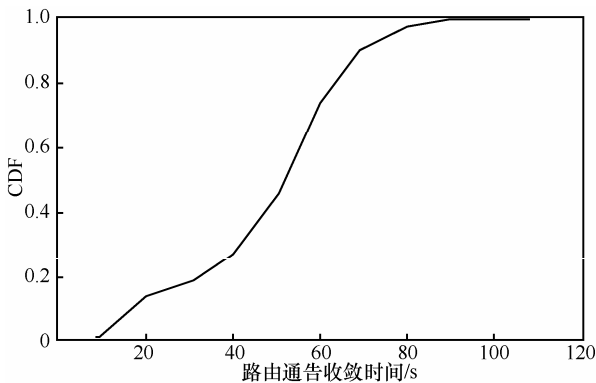


图 8 MIRNDN 下各 AS 路由通告收敛时间 CDF

图 9 给出了各 AS 路由撤销收敛时间的 CDF。需要注意的是，图 9 中关于路由撤销收敛时间的测量结果是仅关于所用 AS 拓扑中上层关系图不存在环路的节点的(从理论上说，从一个 AS 节点到 tier 1 节点是不应存在环路的，而本文实验所用的 AS 级别拓扑是 CAIDA 推测的结果，其中，存在环路，只有小部分节点的上层关系图不存在环路)。从图 9 可以看出，AS 节点路由撤销信息的收敛时间在 5~46 ms，约 60%节点路由撤销的收敛时间小于 10 ms，约 95%节点路由撤销的收敛时间小于 25 ms。MIRNDN 下 AS 节点路由由撤销收敛时间比较短是因为它只传播内容的可达性而没有具体路径的更新

和探索，而且路由撤销的转发始终不受 minrouteadver 控制。此外，该实验结果表明在实际的 AS 级别拓扑中，即使 AS 节点到 tier 1 AS 节点的最长路径也较短，从侧面证实 MIRNDN 运用多条可用路径作为备选路径的可行性。

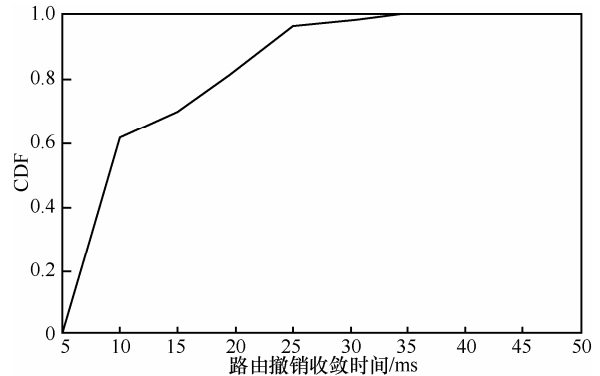


图 9 MIRNDN 下各 AS 路由撤销收敛时间 CDF

图 10 给出了各 AS 路由更新收敛期间发送更新报文数上限的 CDF。从图 10 看到，约 21%节点在路由收敛期间发送的更新报文数小于 22，约 79%节点在路由收敛期间发送的更新报文数上限在 1 000~1 600。对拓扑数据的进一步分析发现，路由更新收敛期间发送更新报文数上限小于 22 的节点的上层关系图比较简单，通往 tier 1 供应者的路径上各节点的供应者 AS 数量都比较少(一般都是 1 或 2，最大的不超过 9)，而且这些节点靠近网络的核心；而路由更新收敛期间发送更新报文数大于 1 000 的节点的上层关系图相对复杂，通往 tier 1 供应者的路径上节点的供应者 AS 数量相对较多(有些节点的供应者数量大至 33)，这些节点主要是根网络节点。该测量结果表明 MIRNDN 下路由更新收敛期间所需通信开销与生成路由更新的节点在 AS 层次图中的位置相关。

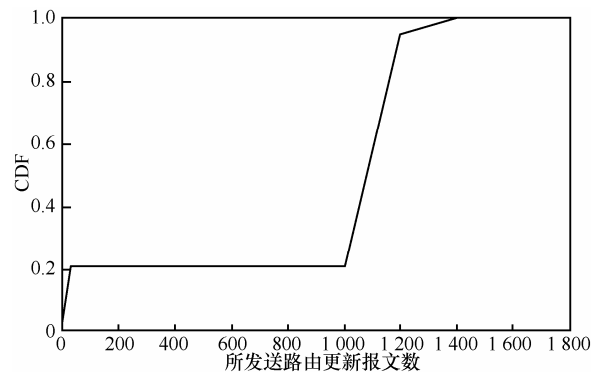


图 10 MIRNDN 下各 AS 路由更新收敛期间发送更新报文数上限的 CDF

### 5.3 数据平面开销

在简单泛洪中, Interest 报文会在每条链路上转发一次, 75 001 条有向边也就意味着任意请求者发送的 Interest 报文都将被转发 75 001 次。在 PPF 中, 节点避免将来自于供应者 AS 和对等 AS 的 Interest 报文向其他供应者 AS 和对等 AS 转发。图 11 显示了 PPF 下每个 AS 节点的(不同)Interest 报文平均被转发次数的 CDF。可以看到 PPF 将 Interest 报文平均被转发的次数减少到 1 601~54 675。因此通过应用“无谷底”路由策略, 即使是最坏的情况, 与简单泛洪相比, 也可以减少  $\frac{75\,001 - 54\,675}{75\,001} \approx$

27% 的 Interest 报文的转发, 而且可以看到近 70% AS Interest 报文平均被转发的次数在 43 850 以下, 相对于简单泛洪大约是  $\frac{75\,001 - 43\,850}{75\,001} \approx 41.5\%$  的减少量。

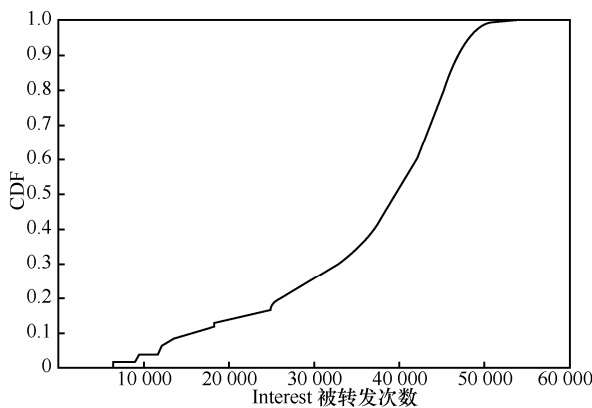


图 11 PPF 下 AS 发送的 Interest 报文平均转发次数的 CDF

为了分析 MIRNDN 中客户 AS 的路由更新进一步过滤不必要 Interest 报文转发的有效性, 本节实验统计了 PPF 下发送的 Interest 报文平均被转发次数最多的 AS, AS 47 885, 在 MIRNDN 下请求不同 AS 内容的 Interest 报文被转发次数的 CDF 如图 12 所示。虽然这不完全代表 MIRNDN 下的最坏情况, 但其为分析提供了一个参考点。图 12 显示 AS 47 885 访问其他 AS 内容的 Interest 报文被转发的次数均在 7 040 左右。这一方面表明 MIRNDN 下客户 AS 的路由更新在 PPF 的基础上进一步过滤了约  $\frac{54\,675 - 7\,040}{54\,675} \approx 87\%$  的 Interest 报文的转发。另一方面表明 AS 47 885 访问其他 AS 内容需要转发 Interest 报文的次数相近; 对 AS 47 885 附近拓扑的分析解

释了这一现象: AS 47 885 有 2 个供应者 AS, 它访问其他 AS 内容所发送的 Interest 报文都向这 2 个供应者 AS 节点转发, 而这 2 个供应者的客户 AS 数量少, AS 47 885 所访问内容绝大多数不在这 2 个供应者 AS 的客户 AS 中, 这 2 个供应者 AS 需将 AS 47 885 请求这些内容所发送的 Interest 进一步向其上层供应者 AS 和对等 AS 转发直到 tier 1 AS, 此时要么找到内容所在客户网络, 要么向其他对等 tier 1 AS 转发。

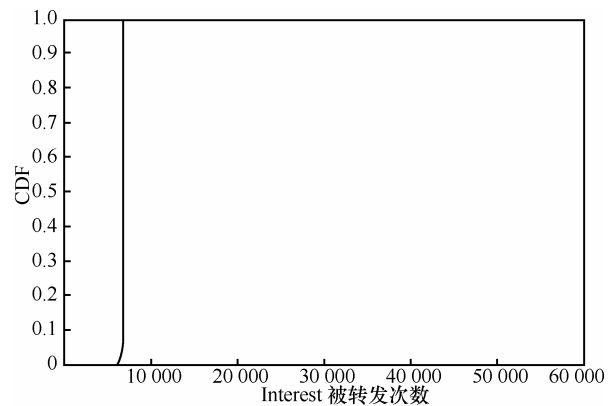


图 12 MIRNDN 下 AS 47885 Interest 报文被转发次数的 CDF

上述仿真结果表明 MIRNDN 的 2 级“过滤器”比较有效地减少了 Interest 报文的不必要转发。值得一提的是, 本节实验假设用户请求的内容只有一个数据源, 在网络中不存在缓存的副本, 而且 NDN 的自适应转发也未启用, 因此上述仿真实验给出的 Interest 报文转发次数的统计结果是最差情况下的。在实际 NDN 网络中, data 报文会缓存在网络中, 当遇到缓存 data 报文的节点时 Interest 报文不会被进一步转发; 另外, 启用自适应转发的网络节点也会选择性地 Interest 仅向存在潜在数据源的接口转发。

### 6 结束语

本文提出了一种 NDN 域间多路径路由机制 MIRNDN。MIRNDN 下 AS 节点依据自身收集的本网络所提供内容的可达性信息、客户 AS 节点向它通告的客户网络中内容的可达性信息以及“无谷底”路由策略原则生成转发表 FIB, 为 Interest 报文的多路径转发提供依据。MIRNDN 是基于内容名字前缀的多路径路由机制, 提高内容获取的健硕性。本文从理论上分析了 MIRNDN 下 FIB 表大小、路由更新的收敛时间和通信开销以及安全特性, 并通过在实际因特网 AS 级别拓扑上仿真 MIRNDN 初步

验证了 MIRNDN 缓解了域间路由的可扩展性问题, 其路由更新的收敛时间较短, 通信开销适量, 并有效地减少了 Interest 报文的不必要转发。MIRNDN 是 NDN 域间路由机制研究的第一步, 下一步需要在 NDN 测试床中运行大量实验, 对在启用网络缓存及自适应转发平面情况下 MIRNDN 的性能进行评估并对 MIRNDN 作进一步优化。

### 参考文献:

- [1] JACOBSON V, SMETTERS D, THORNTON J, *et al.* Networking named content [A]. Proceedings of the 5th International Conference On Emerging Networking Experiments and Technologies [C]. Rome, Italy, 2009. 1–12.
- [2] ZHANG L X, CLAFFY K, CROWLEY P, *et al.* Named data networking [J]. ACM SIGCOMM Computer Communications Review, 2014, 44(3): 66-73.
- [3] JACOBSON V, BURKE J, ESTRIN D, *et al.* Named Data Networking Project Annual Report [R]. Technical Report NDN-0001, 2010.
- [4] CCNx project[EB/OL]. <http://www.ccnx.org>, 2014.
- [5] YI C, AFANASYEV A, WANG L, *et al.* Adaptive forwarding in named data networking [J]. ACM SIGCOMM Computer Communications Review, 2012, 42(3): 62-67.
- [6] DAI H C, LU J Y, WANG Y, *et al.* A two-layer intra-domain routing scheme for named data networking [A]. IEEE Global Communications Conference 2012 [C]. Anaheim, California, USA, 2012. 1-6.
- [7] TORRES J, FERRAZ L, DUARTE O. Controller-based Routing Scheme for Named Data Network [R]. Technical Report, Electrical Engineering Program, COPPE/UFRJ, 2012.
- [8] HOQUE M, AMIN S, ALYYAN A, *et al.* NLSR: named-data link state routing protocol [A]. Proceedings of the 3th ACM SIGCOMM Workshop on Information-centric Networking [C]. Hong Kong, China, 2013. 15-20.
- [9] DIBENEDETTO S, PAPADOPOULOS C, MASSEY D. Routing policies in named data networking [A]. Proceedings of the First ACM SIGCOMM Workshop on Information-centric Networking[C]. Toronto, Ontario, Canada, 2011. 38-43.
- [10] We knew the web was big[EB/OL]. <http://googleblog.blogspot.com/2008/07/we-knew-webwas-big.html>, 2008.
- [11] GAO L. On inferring autonomous system relationships in the Internet [J]. IEEE/ACM Transactions on Networking, 2001, 9(6): 733-745.
- [12] HUSTON G. Interconnection, peering, and settlements[EB/OL]. [https://www.isoc.org/inet99/proceedings/1e/1e\\_1.htm](https://www.isoc.org/inet99/proceedings/1e/1e_1.htm), 1999.
- [13] BGP routing table analysis reports[EB/OL]. <http://bgp.potaroo.net/>, 2014.
- [14] YI C, ABRAHAM J, AFANASYEV A, *et al.* On the role of routing in named data networking [A]. Proceedings of the First ACM Conference on Information-Centric Networking [C]. Paris, France, 2014. 1-10.
- [15] 陈曦. 基于 BGP 协议的 NDN 系统域间内容路由的研究与实现[D]. 北京邮电大学, 2013.  
CHEN X. BGP Based Inter-Domain Content Routing for Named Data Networking[D]. Beijing University of Posts and Telecommunications, 2013.
- [16] PAPADOPOULOS F, KRIOUKOV D, BOGUŃÁ M. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces [A]. The IEEE Conference on Computer Communications 2010 [C]. San Diego, CA, 2010. 2973-2981.
- [17] AFANASYEV A, MAHADEVAN P, MOISEENKO I. Interest flooding attack and countermeasures in named data networking [A]. Proceedings of IFIP Networking 2013 [C]. New York, USA, 2013. 1-9.
- [18] The CAIDA AS relationships dataset[EB/OL]. [http://www.caida.org/data/request\\_user\\_info\\_forms/as\\_relationships.xml](http://www.caida.org/data/request_user_info_forms/as_relationships.xml), 2010.
- [19] DIMITROPOULOS X, KRIOUKOV D, FOMENKOV M. As relationships: inference and validation [J]. ACM SIGCOMM Computer Communications Review, 2007, 37(1): 29-40.

### 作者简介:



胡晓艳 (1985-), 女, 江西临川人, 博士, 东南大学讲师, 主要研究方向为信息中心网络和网络管理。



龚俭 (1957-), 男, 上海人, 博士, 东南大学教授、博士生导师, 主要研究方向为网络体系结构、网络入侵检测和网络管理。