

内持续不可用,则需要将其分解为之前被聚合的路由,以提升路由准确性,改善路由可达情况.为了支撑聚合路由回滚,需要在各个节点保留如图 4 所示的线索树结构,在聚合路由的同时保留被聚合路由信息,以便更新和回滚聚合路由.其中,当前路由线索是当前使用的聚合路由的线索,当需要更新和回滚路由时,可以快速找到相应的聚合路由,并回滚或更新相关被聚合路由.本文的动态聚合机制可以在保证路由可用性的前提下最大程度地聚合路由,缩小全网路由规模,提升层次化名字路由效率.

4 实验评估

4.1 方案部署与实现

我们在 NDN 网络的仿真平台 ndnSIM^[32]的基础上实现了本文针对层次化名字路由的聚合机制.

- 首先,我们对 ndnSIM 的 FIB 模块进行了修改.现有的 ndnSIM 中,每条路由包含路由名字前缀和转发接口信息.基于本文动态路由聚合机制,我们在转发接口信息中添加了后缀压缩表示 C 和可达性统计 RP(见公式(9)).RP 中的统计值 $n(t-1)$ 是在查表转发请求时计数,而 $n_r(t)$ 是下一个时刻收到的内容包的数量.根据不同的预定容量 n ,用于压缩表示后缀名字段的 CCBF 按布隆过滤器最优配置设置比特数组长度 m 及哈希函数个数 k ^[26],同时,根据本文第 3.2 节的分析,比特数组数量被设置为 16.在此基础上,FIB 模块周期性更新 RP,根据更新过的 RP 分析现有路由的可用性,将持续可用的路由进行进一步聚合,得到新的聚合路由,并周期性地通告周边节点.同时更新在低速外部存储中的记录路由聚合过程的数据结构(如图 4 所示),以备聚合路由回滚时还原其中备份的路由信息.
- 其次,我们修改了转发模块 ndn-forwarding-strategy.实现了基于名字前缀和后缀压缩表示的转发机制.

仿真平台在本地计算机上部署,其配置如下:Ubuntu 13.04,内核版本 3.8.8,Intel Core i7 3.4G CPU,DDR3 1866 16G 内存,1TB 硬盘.为了保证仿真结果真实、可靠,我们采用了一个真实的网络拓扑——欧洲 EBONE 网络拓扑^[33],包括 279 个节点、731 条边,其中包括 65 个边界网关节点、45 个网关节点和 169 个边缘路由节点.在边缘节点中,我们随机选择了 10 个节点作为内容发布节点 producer,实验中请求的各类内容全部由这 10 个节点发布.另外,选择了 30 个节点作为内容请求方 consumer,内容请求服从参数 α 为 0.6 的 Zipf-Mandelbrot 分布,即被请求的多数内容为热门内容,其他内容很少被请求.为了全面验证本文路由聚合机制的有效性,实验中使用了多个不同规模的数量集,这些数据集的具体信息见表 1.

Table 1 Datasets

表 1 数据集

数据集名称	squidGuard blacklist	Shalla's blacklists	Fabrice Prigent's blacklist	MESD blacklists	Alexa top 1M
大小(KB)	27 796	9 791	8 449	8 457	9 787
条目数	4 583 626	1 700 000	1 682 512	1 342 541	1 000 000

在此基础上,通过对这些数据进行组合,我们可以得到各种规模的名字数据集.在其基础上,我们对本文路由聚合机制的有效性进行了全面的评估.首先比较了路由聚合前后的路由表规模;其次,通过与 NDN 原生路由查表机制以及利用 CBF 压缩路由表的对比方案比较,给出了路由聚合前后的路由查表效率的比较;最后,通过对比分析了本文路由聚合机制带来的假阳性问题.

4.2 聚合前后路由表规模对比分析

路由聚合的目标是减少路由条目数量,从而提升路由查表效率,进而优化网络传输效率.同时,路由条目的减少还会减少路由设备内存的使用,降低 NDN 网络部署成本.为了验证本文路由聚合机制的聚合效果,我们在数据集 Alexa top 1M 上比较了各个路由节点在同一时刻采用聚合后的路由条目数量和不采用聚合的情况下路由条目数量的比值 R ,在 2 000s 内对 279 个路由器每隔 6s 记录一次 R 值,仿真结果如图 5 所示.

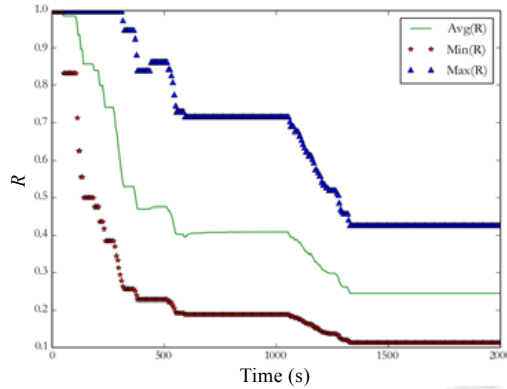


Fig.5 Effecton on FIB size reduction of route aggregation

图 5 路由聚合对路由表规模的影响

仿真过程中,随着时间 Time 的推移,路由聚合程度不断提升,路由规模不断缩小,在 1400s 附近,全网路由聚合达到稳定状态.全部 279 个节点中,在全网路由聚合完成后,路由条目数量至少压缩了 60%,平均压缩了 75.47%,路由聚合对路由规模的缩减效果明显.注意,由于现有数据集中多数名字(URL)只有 2 级,如果应用于实际部署的 NDN 网络中,路由名字层次变化更多,路由聚合效果也会更为显著.

4.3 聚合路由效率分析

作为影响网络传输效果的关键因素,路由查表效率直接影响网络数据包转发延时.路由聚合主要通过减少路由条目数量(路由表规模)来提升路由查表效率,进而缩小网络数据包转发延时.利用本文路由聚合机制,可以将多条具有公共前缀的层次化名字路由聚合为单条聚合路由,并利用这些路由的公共名字前缀和后缀压缩表示来共同标识该聚合路由.因此,与未进行聚合的层次化名字路由相比,本文的路由聚合机制在缩短名字前缀查找时间的同时,也引入了路由后缀的查找过程.为了验证本文聚合过程对路由查表时间的影响,我们首先针对不同规模 N_s 的名字路由(1 万、10 万和 100 万)进行路由聚合,并比较聚合前后各路由的查表时间总和和 LTime.实验结果如图 6(a)所示,圆形点线代表未聚合路由查表时间总和,方形点线为本文聚合路由查表时间总和.可以看出,未被聚合的路由的查表时间与路由规模成正比例关系,随着路由规模 N_s 的增长线性增加;而聚合后路由查找基本不受路由规模增长的影响,仅表现出 1s 的差异,如图 6(a)所示.究其原因,主要是由于聚合后减小了前缀数量及前缀查找时间;同时,本文在后缀压缩表示查询方面的优化(为 CCBF 添加 16 个比特数组按位或得到的归并结果 *orBarr*)也避免了大量后缀查询时延的引入.

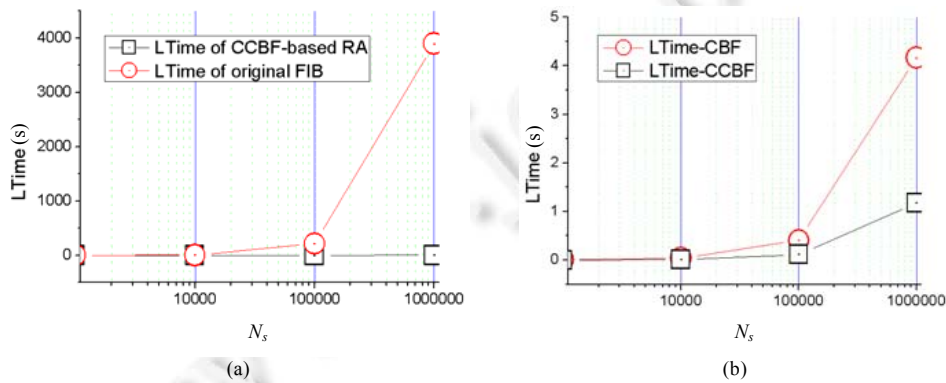


Fig.6 Comparison of route lookup time

图 6 路由查找效率对比

为了进一步验证本文路由聚合机制的查表效率,我们在 CBF 的基础上构建了一套路由压缩机制作为对比方案,采用与本文的聚合路由相同的前缀和 CBF 压缩表示的路由后缀共同标识压缩后的路由.压缩后的路由规模与本文路由聚合后的路由规模相同,通过这一方式对比验证本文聚合路由查表效率.如图 6(b)所示,方形点线为 CBF 的查询时间 LTime-CBF,圆形点线为本文 CCBF 的查询时间 LTime-CCBF.在不同规模的路由表上,本文方案的查表时间始终只有对比方案的一半,本文建立在 *orBarr* 上的查询过程的效率要优于 CBF 的查询效率.

同时,我们对比了聚合前后路由更新效率(如图 7 所示),路由更新时间等于所有路由的查找、删除和添加操作的总和,用 UpTime 代表.图 7(a)中,圆形点线为未聚合路由总更新时间,方形点线为本文方案路由总更新时间.未聚合路由的整体更新时间都随路由规模 N_s 的增长呈线性增长,本文路由由聚合机制的更新时间增加较慢,路由规模对路由更新效率的影响同样存在.为了进一步评估本文机制的路由更新效率,图 7(b)中,我们对比了本文机制的路由更新时间和基于 CBF 的对比方案的路由更新时间,圆形点线为本文机制的总路由更新时间,方形点线为对比方案的总路由更新时间.两个方案的总路由更新时间都随路由规模 N_s 的增长呈线性增长,本文路由更新时间略小于对比方案路由更新时间的 2 倍,虽然与路由查表时间相比,本文的路由聚合机制的路由更新时间较长,但路由更新不是网络路由转发过程中最关键的的操作,操作频率较低,且独立处理,基本不影响路由查表过程.

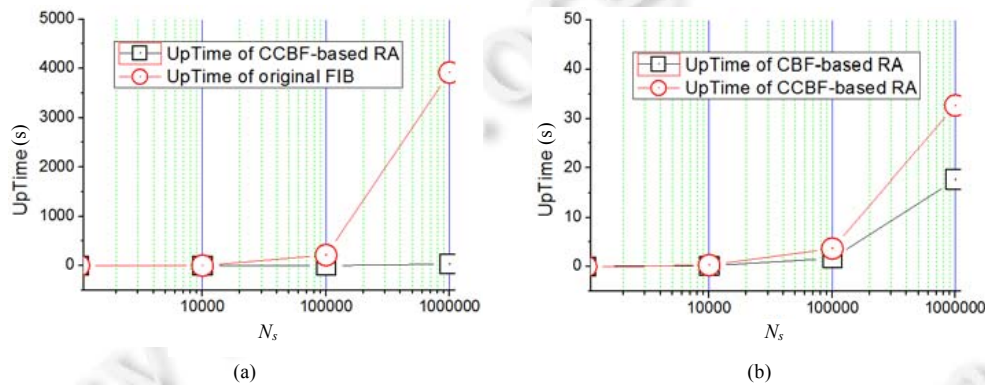


Fig.7 Comparison of route updating time

图 7 路由更新时间对比

4.4 假阳性分析

本文的聚合标识在使用路由名字前缀的同时,还引入了基于 CCBF 的后缀压缩表示.CCBF 本身可能会在查询过程中引入假阳性,即查询到未被压缩表示过的名字段.假阳性在路由聚合问题的背景下,将会引发路由查询的假阳性误判,将数据包发往没有通告过相关路由的接口.注意,假阳性不会影响数据包从通告过该路由的接口正常转发,仅仅是增加了冗余的路由转发.本节实验验证了本文后缀压缩表示所采用的 CCBF 布隆过滤器的假阳性情况,并与基于 CBF 的对比方案的假阳性进行了对比.在布隆过滤器的配置过程中,本文为各个过滤器设定的最大假阳性率均为 0.01.图 8 中,菱形点线为 CCBF 在不同规模数据集下的查询假阳性数 Fp -CCBF,方形点线为对比方案在不同规模数据集下的查询假阳性数 Fp -CBF.两种布隆过滤器在不同规模的数据集上的假阳性值相同.

为了进一步分析假阳性的出现情况,分析两种方案在不同的添加比例 R_{fill} 下的假阳性变化情况,本文针对 3 种不同规模的数据集(1 万、10 万和 100 万)进行了实验.其中, $R_{fill}=n_c/n$, n 为布隆过滤器容量(分别取 1 万、10 万和 100 万), n_c 为添加到布隆过滤器的名字段数量.如图 9 所示,在不同规模的数据集下,两种方案的假阳性查询数量在 R_{fill} 达到 0.85 后出现了假阳性查询,并在 R_{fill} 到达 1 时达到最大值,两种方案的假阳性查询数量始终保持相同.本文方案在查询假阳性方面等同于基于计数布隆过滤器 CBF 的对比方案,没有带来额外的假阳性查询.

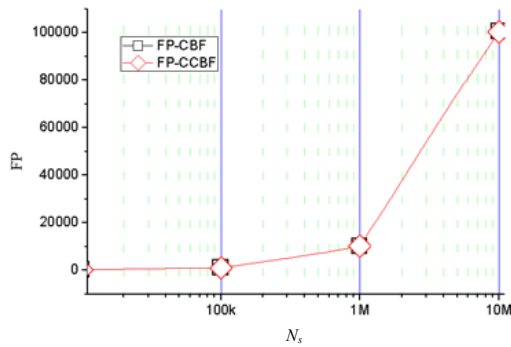
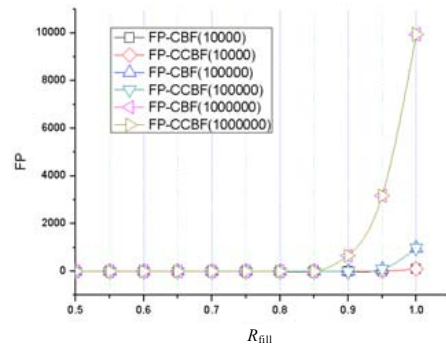


Fig.8 False positive under datasets with different size

图8 不同规模数据集下的假阳性

Fig.9 False positive with different R_{fill} 图9 不同 R_{fill} 下的假阳性

5 结论

为了实现针对层次化名字路由的聚合机制,进而优化NDN等信息中心网络的传输效率,本文首先分析了层次化名字路由的聚合问题,明确了路由聚合过程中需要同时利用路由名字前缀和后缀压缩表示来共同标识聚合后的路由.为了实现名字后缀的压缩标识,同时在路由聚合过程中通告和进一步聚合(合并)这些后缀压缩标识,本文提出了一种全新的计数布隆过滤器 CCBF.CCBF 在支持多过滤器合并的同时,还优化了查询效率.在此基础上,本文进一步提出了面向路由可达性的动态路由聚合机制,在保证路由可达性的前提下最大化聚合路由.我们在真实网络拓扑及数据集上构建了实验环境,验证了本文路由聚合机制的有效性,本文路由聚合机制通过有效聚合路由,以较小的冗余路由转发(假阳性路由查询)为代价,减小了路由规模,优化了网络传输效率,为NDN等信息中心网络投入实际应用提供了前提.

本文的路由聚合程度由路由可达性动态决定.作为影响可达性的关键因素,如果能够有效地降低 CCBF 的假阳性查询结果,则不但可以避免过多的冗余转发,而且可以提升路由可达性,进一步提升路由聚合程度,减少路由条目数量.为此,在下一步的研究工作中,应该将综合分析现有计数布隆过滤器的优化思路,降低堆叠计数布隆过滤器的查询假阳性,进一步改进路由后缀压缩表示机制及相应的路由聚合过程.

References:

- [1] Bari MF, Chowdhury SR, Ahmed R, Boutaba R, Mathieu B. A survey of naming and routing in information-centric networks. IEEE Communications Magazine, 2012,50(12):44–53. [doi: 10.1109/MCOM.2012.6384450]
- [2] Koponen T, Chawla M, Chun BG, Ermolinskiy A, Kim KH, Shenker S, Stoica I. A data-oriented (and beyond) network architecture. In: Proc. of the 2007 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2007. 181–192. [doi: 10.1145/1282427.1282402]
- [3] Roberts J. The clean-slate approach to future internet design: A survey of research initiatives. Annals of Telecommunications, 2009, 64(5):271–276. [doi: 10.1007/s12243-009-0109-y]
- [4] Xie GG, Zhang YJ, Li ZY, Sun Y, Xie YK, Li ZC, Liu YJ. A survey on future internet architecture. Chinese Journal of Computers, 2012,35(6):1109–1119 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2012.01109]
- [5] Wu C, Zhang YX, Zhou YZ, Fu XM. A survey for the development of information-centric networking. Chinese Journal of Computers, 2015,38(3):455–471 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2015.00455]
- [6] Jacobson V, Smetters DK, Thornton JD, Plass MF, Briggs NH, Braynard RL. Networking named content. In: Proc. of the 5th Int'l Conf. on Emerging Networking Experiments and Technologies. New York: ACM Press, 2009. 1–12. [doi: 10.1145/1658939.1658941]
- [7] Zhang LX, Jacobson V, Zhang BC, Tsudik G, Claffy KC, Massey D, Abdelzaker T, Wang L, Crowley P, Yeh E. Named data networking (NDN) project. Technical Report, NDN-0001, Los Angeles: UCLA, 2017.

- [8] Wu H, Shi JX, Wang YX, Wang YL, Zhang G, Wang Y, Liu B, Zhang BC. On incremental deployment of named data networking in local area networks. In: Proc. of the Symp. on Architectures for Networking and Communications Systems. Piscataway: IEEE Press, 2017. 82–94. [doi: 10.1109/ANCS.2017.18]
- [9] Liang CC, Yu FR, Zhang X. Information-centric network function virtualization over 5G mobile wireless networks. IEEE Network, 2015,29(3):68–74. [doi: 10.1109/MNET.2015.7113228]
- [10] Wang Y, He KQ, Dai HC, Meng W, Jiang JC, Liu B, Yan C. Scalable name lookup in ndn using effective name component encoding. In: Proc. of the 2012 IEEE 32nd Int'l Conf. on Distributed Computing Systems. Piscataway: IEEE, 2012. 688–697. [doi: 10.1109/ICDCS.2012.35]
- [11] Wang Y, Zu Y, Zhang T, Peng KY, Dong QF, Liu B, Meng W, Dai HC, Tian X, Xu ZH, Wu H, Yang D. Wire speed name lookup: A GPU-based approach. In: Proc. of the 10th USENIX Conf. on Networked Systems Design and Implementation. Berkeley: USENIX Association, 2013. 199–212.
- [12] Shi JX, Liang T, Wu H, Liu B, Zhang BC. NDN-NIC: Name-based filtering on network interface card. In: Proc. of the 3rd ACM Conf. on Information-centric Networking. New York: ACM Press, 2016. 40–49. [doi: 10.1145/2984356.2984358]
- [13] So W, Narayanan A, Oran D, Wang YG. Toward fast NDN software forwarding lookup engine based on hash tables. In: Proc. of the 2012 ACM/IEEE Symp. on Architectures for Networking and Communications Systems. Piscataway: IEEE, 2012. 85–86. [doi: 10.1145/2396556.2396575]
- [14] So W, Narayanan A, Oran D, Stapp M. Named data networking on a router: Forwarding at 20Gbps and beyond. In: Proc. of the ACM SIGCOMM 2013 Conf. on SIGCOMM. New York: ACM Press, 2013. 495–496. [doi: 10.1145/2486001.2491699]
- [15] Le F, Xie GG, Zhang H. On route aggregation. In: Proc. of the 7th Conf. on Emerging Networking Experiments and Technologies. New York: ACM Press, 2011. 1–12. [doi: 10.1145/2079296.2079302]
- [16] Degermark M, Brodnik A, Carlsson S, Pink S. Small forwarding tables for fast routing lookups. In: Proc. of the ACM SIGCOMM'97 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication. New York: ACM Press, 1997. 3–14. [doi: 10.1145/263105.263133]
- [17] Dharmapurikar S, Krishnamurthy P, Taylor DE. Longest prefix matching using Bloom filters. In: Proc. of the 2003 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2003. 201–212. [doi: 10.1145/863955.863979]
- [18] Pong F, Tzeng NF. Concise lookup tables for IPv4 and IPv6 longest prefix matching in scalable routers. IEEE Trans. on Networking, 2012,20(3):729–741. [doi: 10.1109/TNET.2011.2167158]
- [19] Yi C, Afanasyev A, Wang L, Zhang BC, Zhang LX. Adaptive forwarding in named data networking. ACM SIGCOMM Computer Communication Review, 2012,42(3):62–67. [doi: 10.1145/2317307.2317319]
- [20] Wang L, Hoque AKM, Yi C, Alyyan A, Zhang BC. OSPFN: An OSPF based routing protocol for named data networking. Technical Report, NDN-0003. Memphis: University of Memphis, 2012.
- [21] Hoque AKMM, Amin SO, Alyyan A, Zhang BC, Zhang LX, Wang L. NLSR: Named-data link state routing protocol. In: Proc. of the 3rd ACM SIGCOMM Workshop on Information-centric Networking. New York: ACM Press, 2013. 15–20. [doi: 10.1145/2491224.2491231]
- [22] Li F, Chen FY, Wu JM, Xie HY. Fast longest prefix name lookup for content-centric network forwarding. In: Proc. of the 2012 ACM/IEEE Symp. on Architectures for Networking and Communications Systems. Piscataway: IEEE, 2012. 73–74. [doi: 10.1145/2396556.2396569]
- [23] Dai HC, Lu JY, Wang Y, Liu B. A two-layer intra-domain routing scheme for named data networking. In: Proc. of the 2012 IEEE Global Communications Conf. Piscataway: IEEE, 2012. 2815–2820. [doi: 10.1109/GLOCOM.2012.6503543]
- [24] Garcia-Luna-Aceves JJ. A more scalable approach to content centric networking. In: Proc. of the 2015 24th Int'l Conf. on Computer Communication and Networks. Piscataway: IEEE, 2015. 1–8. [doi: 10.1109/ICCCN.2015.7288363]
- [25] Yuan HW, Crowley P. Reliably scalable name prefix lookup. In: Proc. of the 2015 ACM/IEEE Symp. on Architectures for Networking and Communications Systems. Piscataway: IEEE, 2015. 111–121. [doi: 10.1109/ANCS.2015.7110125]
- [26] Broder A, Mitzenmacher M. Network applications of Bloom filters: A survey. Internet Mathematics, 2004,1(4):485–509. [doi: 10.1080/15427951.2004.10129096]

- [27] Fan L, Cao P, Almeida J, Broder AZ. Summary cache: A scalable wide-area Web cache sharing protocol. ACM SIGCOMM Computer Communication Review, 1998,28(4):254–265. [doi: 10.1145/285243.285287]
- [28] Cohen S, Matias Y. Spectral Bloom filters. In: Proc. of the 2003 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2003. 241–252. [doi: 10.1145/872757.872787]
- [29] Bonomi F, Mitzenmacher M, Panigrahy R, Singh S, Varghese G. An improved construction for counting Bloom filters. In: Proc. of the 14th Conf. on Annual European Symp. London: Springer-Verlag, 2006. 684–695. [doi: 10.1007/11841036_61]
- [30] Huang K, Zhang J, Zhang DF, Xie GG, Salamatian K, Liu AX, Li W. A multi-partitioning approach to building fast and accurate counting Bloom filters. In: Proc. of the 2013 IEEE 27th Int'l Symp. on Parallel and Distributed Processing. Piscataway: IEEE, 2013. 1159–1170. [doi: 10.1109/IPDPS.2013.51]
- [31] Béla B. Random Graphs. 2nd ed., Cambridge: Cambridge University Press, 2001. 130–159.
- [32] Afanasyev A, Moiseenko I, Zhang LX. NDN-SIM: NDN simulator for ns-3. Technical Report, NDN-0005. Los Angeles: UCLA, 2012.
- [33] Spring N, Mahajan R, Wetherall D, Anderson T. Measuring ISP topologies with rocketfuel. ACM SIGCOMM Computer Communication Review, 2002,32(4):133–145. [doi: 10.1145/964725.633039]

附中文参考文献:

- [4] 谢高岗,张玉军,李振宇,孙毅,谢应科,李忠诚,刘韵洁. 未来互联网体系结构研究综述. 计算机学报, 2012,35(6):1109–1119. [doi: 10.3724/SP.J.1016.2012.01109]
- [5] 吴超,张尧学,周悦芝,傅晓明. 信息中心网络发展研究综述. 计算机学报, 2015,38(3):455–471. [doi: 10.3724/SP.J.1016.2015.00455]



许志伟(1979—),男,内蒙古呼和浩特人,博士,讲师,CCF 专业会员,主要研究领域为互联网传输,可靠性和安全性优化.



张玉军(1976—),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为未来互联网体系结构.



陈波(1982—),男,博士,助理教授,主要研究领域为互联网传输,可靠性和安全性优化.