维普资讯 http://www.cqvip.com

树链混合组播源认证协议*)

王卫东　李之棠
(华中科技大学计算机学院　武汉 430074)

摘　要　源认证是组播通讯面临的一个挑战性问题,必须为大量接受者提供系统开销低、可靠性高的确认数据来源的方法。本文提出了一种有效的组播源认证协议 HTC,该方案结合 Hash 树和多 Hash 链方法的优点,有效地降低了通讯开销。采用二态马尔科夫丢包模型进行了大量的仿真实验,获得了一个最优的 Hash 跨度组合 1-2-7-11-16-20-25-30。与已有多个认证方案进行比较,说明 HTC 是一种有效的组播源认证方案。

关键词　组播,源认证,树链混合

Tree and Chains United into Multicast Authentication

WANG Wei-dong　LI Zhi-tang
(College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract　For multicast communication, authentication is a challenging problem, since it requires that a large number of recipients must verify the data originator. We propose an efficient multicast source authentication protocol called Hybrid Tree and Chains scheme (HTC), which shows more loss resistibility, less communication cost. The HTC scheme is based on combination of single Chain scheme and Hash Tree Chains scheme, and integrates the advantages of both. In this scheme, stream is firstly divided into blocks with n packets, and each block consists of m clusters, everyone of which contains a tree of packets. All clusters are chained together. Through HTC, packets of one cluster can be authenticated by any packet of the previous cluster. Compared to other multicast authentication protocols, the proposed scheme has the following advantages: 1) dramatically improves the resistance to burst packets loss, 2) low computation and communication overhead.

Keywords　Multicast, Source authentication, Hybrid tree and chains

1　引言

近年来,电话会议、付费媒体收看、股票实时信息发布等具有明显组播通讯特点的应用越来越广泛,组播通讯安全越来越受到重视。机密性、数据来源认证、完整性认证和抗否认性是组播通讯的安全要求。机密性是保证未授权用户无法获取数据内容;数据来源认证需要保证接受的数据不是他人冒名发送的;完整性认证保证数据的篡改可以被用户发现;抗否认性则保证发送方无法否认自己曾经发送的数据。

不同的组播应用,有不同的安全需求。在很多应用中传输的数据并不需要保密,但是数据来源的确认是必需的。对于组播通讯而言,提供源认证是富有挑战性的工作,因为组播通讯的接受者可能数量巨大,网络环境各异。采用类似单播通讯源认证的方法是为每个接受者维护一个共享密钥,发送方对发送的信息为每个人计算消息认证码(Message Authentication Codes),并把这些 MACs 附在组通信消息发送给接受者。对于规模较大的组播应用,这将带来极高的计算开销(发送方)和组通讯开销。基于公私钥体系的签名算法可以大大降低系统开销。在公私钥体系中,接受者可以提前以其他途径获得发送者的公钥,例如邮件。发送者只需要对自己发送的整个报文的哈希值采用私钥签名(加密),接收方采用发送者的公钥对收到的签名信息解密,并比较信息的哈希值,即可确定数据来源是否可信。由于公私钥体系的计算开销过高,对每个消息都采用签名(发送方)、解密-哈希比较(接收方),对于通信量较大的组播应用,如流媒体服务,一般用户是无法承受的。解决问题的途径有两个:一是提高公私钥加解密算法的效率、降低计算量,另一个是采用签名开销分担策略[1,2]。

文献[1]最早提出了采用单哈希链结构来分担签名开销的思路,发送方和每一个接收方各自只需要分别做一次加密(签名)和解密(签名验证)计算就可以实现对很多报文的来源认证,大大减少了计算开销。同时,由于签名信息较长,分担策略同时还减轻了通信开销。

无论是 IP 层组播还是应用层组播,由于 TCP/IP 的基础是尽力转发,报文丢失无法避免。任何报文的丢失都可能影响单哈希链的连续性,导致对报文的验证失败。因此认证信息必须提供一定的冗余,这样即使部分报文丢失,认证信息也可以得到恢复,以认证其它报文。另外,文献[3]的研究表明,Internet 上报文的丢失具有很强的突发性,即当一个报文丢失,其后续报文的丢失概率会增加,这为冗余认证方案的研究带来了困难。为了解决这种突发丢包现象,文献[4-9]分别提出了不同的冗余认证方案。与其他方案相比,我们的认证方案针对突发性报文丢失特征,结合树链混合结构的优势,以很小的发送方延迟换来了认证效率的显著提高。

*)本文得到国家自然科学基金项目"P2P 网络的关键安全问题研究"(60573120)资助。王卫东　博士生,主要研究方向为 P2P 网络安全、计算机网络及安全;李之棠　教授,博士生导师,博士,主要研究方向为网络与信息安全、光互连与光计算。

维普资讯 http://www.cqvip.com

在第2节，详细介绍HTC的认证处理过程；第3节讨论HTC的分析和仿真结果；第4节详细列出HTC与已有的几种相关协议的比较，最后是全文的总结。

2 混合树链机制(HTC)

2.1 认证方法介绍

HTC认证方案的具体签名步骤描述如下：

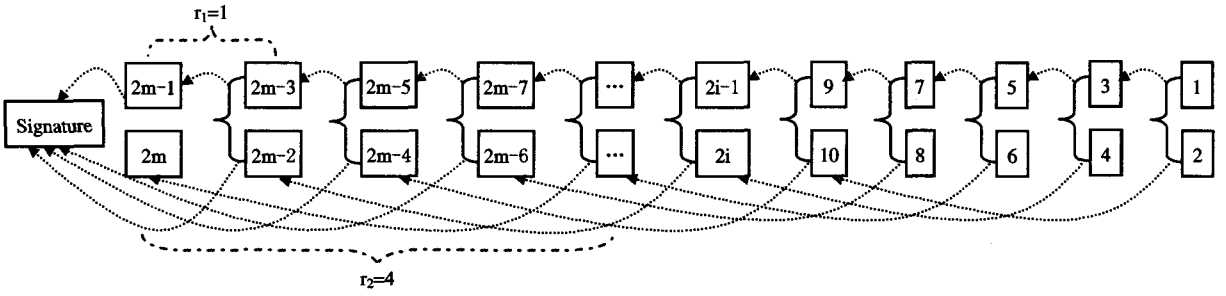
定义1 $\|$为串接符。报文流首先被分割为若干块，即$B=(B_1 \| B_2 \| \cdots)$；每个块包括n个报文，即$B_i=(M_{(i-1)n} \| \cdots \| M_{in-1})$，这$n$个报文分担一个签名开销，所有的块的认证处理相同。

定义2 每一块的n个报文被分为m条，本文称之为群，即$B_1=(C_1 \| C_2 \| \cdots C_m)$。群的大小一般是2的乘幂，即$m=n/c$，$c=2^k$，$k\in(1,2,\cdots,\log_2 n)$，每一个群包括$c$个连续报文，即$C_i=(M_1^{(i)} \| M_2^{(i)} \| \cdots \| M_c^{(i)})$。

定义3 跨度集合$S=\{r_j, 0<j\leqslant c\}$，其中$r_j$被称之为一个跨度，存在$0<r_j\leqslant m-1$，$R$是$c$个跨度的升序集合，集合内元素可以相等。

定义4 $D^{(i)}$被定义为群C_i的群摘要。同群中的每一个报文隐式携带本群摘要，同时显式携带它群摘要，不同的报文可以携带不同的它群摘要。$P_j^{(i)}$是群C_i的第j个报文，则携带的它群摘要为$D^{(i-r_j)}$，$j\in(1,\cdots,c)$。跨度r_j表示群内第j个报文携带它群摘要的偏移量。

图1表述了群间的认证依赖关系，其中$c=2$；$n=2m$；$R=\{1,4\}$。图中每两个连续报文组成一个群，群摘要由h函数(即Hash函数)通过步骤5)计算。如图所示，对于跨度$r_1=1$，$r_2=4$，则第1群的群摘要由第2群的第1个报文(报文3)和第5群的第2个报文(报文10)携带。对于第5群，除了第2个报文(报文10)携带第1群的群摘要，该群的第1个报文(报文9)还携带第4群的群摘要。最后n_s个群摘要由签名报文携带，存在$n_s=r_c$。



$c=2, n=2m, r_1=1, r_2=4$

图1 构建群间摘要链

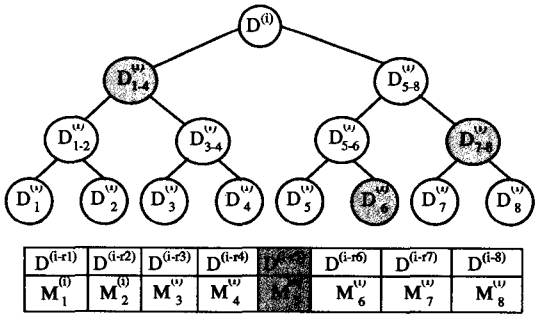
每一块的群摘要$D^{(i)}$按顺序(即$D^{(1)}, D^{(2)}, D^{(3)}, \cdots, D^{(m)}$)计算，图2说明了群中每一个报文需要携带的摘要的计算过程，其中$k=3$，$c=8$，即群的大小为8。详细的计算步骤如下：

a) $D_j^{(i)}=\begin{cases} h(M_j^{(i)}), & i=0 \\ h(M_j^{(i)}) \| D^{(i-r_j)}, & i-r_j\geqslant 0 \end{cases}$

b) $D_{1\text{-}2}^{(i)}=h(D_1^{(i)} \| D_2^{(i)})$，$D_{3\text{-}4}^{(i)}=h(D_3^{(i)} \| D_4^{(i)})$，$\cdots$

c) $D_{1\text{-}4}^{(i)}=h(D_{1\text{-}2}^{(i)} \| D_{3\text{-}4}^{(i)})$，$D_{5\text{-}8}^{(i)}=h(D_{5\text{-}6}^{(i)} \| D_{7\text{-}8}^{(i)})$

d) $D^{(i)}=D_{1\text{-}8}^{(i)}=h(D_{1\text{-}4}^{(i)} \| D_{5\text{-}8}^{(i)})$



$k=3, c=8$

图2 计算群摘要

每一个报文$P_j^{(i)}$(即群C_i的第j个报文)需要携带$1+k$个Hash，其中一个用于携带它群摘要，另外k个Hash携带本群的关联中间Hash。如图2所示，例如$P_5^{(i)}$将携带如下信息：$M_5^{(i)}$，$D_6^{(i)}$，$D_{7\text{-}8}^{(i)}$，$D_{1\text{-}4}^{(i)}$和$D^{(i-r5)}$。

当第一群的所有报文计算完群内Hash后，就可以发送出去，不会影响整个树链的哈希嵌套生成。整个发送端的延迟只是增加了一个群的时间。

与其他多链方案类似，本文假定块的签名信息总能被接收到，这是每一块的报文可以被认证的基础。接收端收到报文后，缓冲等待签名报文的到来。签名包到达后，公钥解密得到块摘要，并比较签名包中携带的群摘要列表的哈希结果，确定签名包可信且没有被篡改。由于$D^{(m)}$包含在签名信息中，$D^{(m)}$可信，群C_m的报文可以立即得到认证；$P_j^{(m)}$得到认证，其携带的它群摘要$D^{(1-r_j)}$也是可信的，根据$D^{(1-r_j)}$就可以认证其他群的报文。为了方便算法的说明和仿真，我们引入Hash缓冲池(HB)，具体步骤如下(假定报文$P_5^{(i)}$)：

a) 计算$D'^{(i)}=h(D_{1\text{-}4}^{(i)} \| h(h(hM_5^{(i)}) \| D^{(i-r_5)}) \| D_6^{(i)} \| D_{7\text{-}8}^{(i)})$

b)在Hash缓冲池中查找并比较$D^{(i)}=D'^{(i)}$。如果相等，该报文认证通过；

c) 查看$P_5^{(i)}$携带的$D^{(i-r_5)}$是否已存入HB，如果没有，将$D^{(i-r_5)}$加入HB。

2.2 认证概率

可以推导出无突发丢包情况时的验证成功率。

定义5 报文验证成功条件概率$Pr_j^{(i)}=Pr$(报文$P_j^{(i)}$是可验证的$|$ $P_j^{(i)}$被接受到)

同群报文依赖相同的群摘要，所以同群报文的认证概率是相同的，存在$Pr_1^{(i)}=Pr_j^{(i)}=\cdots=Pr_c^{(i)}$

定义6 群验证成功条件概率$Pr^{(i)}=Pr$(群C_i的报文是可验证的$|$群C_i的报文被接受)

存在群认证概率$Pr^{(i)}=Pr_j^{(i)}$，$j\in(1,2,\cdots,c)$。

维普资讯 http://www.cqvip.com

定义 7 群哈希关联报文集合 $AS_{(i)}=\{P_j^{(i+r_j)},\delta P_s\}$，其中 $i+r_j\leqslant m$，P_s 为签名包文，

$$\delta=\begin{cases}1,\text{签名包携带该群摘要}\\0,\text{签名包没有携带该群摘要}\end{cases}$$

只有当 $AS_{(i)}$ 中的所有报文都丢失或无法被验证的情况下，$P_j^{(i)}$ 才认证失败。定义 q 为报文丢失概率。对于签名包文直接认证的群，报文验证成功的条件概率为 1，其他报文可以通过公式(1)计算：

$$\Pr_j^{(i)}=\Pr^{(i)}=1-\prod_{k=1}^{c}(1-\Pr^{(i+r_k)}(1-q)) \quad (1)$$

2.3 抗丢包认证分析

报文的认证概率与丢包模型的如下两个参数密切相关：最大突发丢包个数(β)和长程丢包率(π_0)。在其它多链认证方案中，报文携带的每一个 Hash 值只能认证一个报文；在 HTC 方案中，每个报文携带的 k 个 Hash 可以认证 $2^{(k-1)}$ 个报文。另外，通过合理选择 R 中的跨度(r_j)组合，可进一步增加对报文突发性的抵抗能力。

3 仿真与分析

3.1 丢包模型

认证概率与丢包模型直接相关。两种不同的丢失模型用在仿真模拟中：一种是随机丢包模型 RLM，另一种是文献[8]提出的二态马尔科夫链丢失模型(2-state Markov Chain loss model)，该模型将丢包过程建模为两态的马尔科夫随机过程(0 和 1 分别代表丢失和不丢失)，两个长程概率分别是 π_0 和 $\pi_1=1-\pi_0$。为表征突发程度，该模型引入突发丢包平均长度 β，β 越大，表征报文丢失突发性越强。四个转移概率可以表示为 π_1，π_0 和 β 的函数：

$$p_{01}=\frac{1}{\beta},\ p_{10}=\frac{\pi_0}{\beta\pi_1},\ p_{00}=1-\frac{1}{\beta},\ p_{11}=1-\frac{\pi_0}{\beta\pi_1}$$

当 $\beta=1/\pi_1$ 时，等效于随机丢包模型。

3.2 仿真

编写了一个仿真程序，用于测试实际的认证效果。仿真主要在于比较不同 Hash 跨度组合的优劣。每一种测试至少重复 2000 次，测试的详细参数如下。

1) 块的大小(n)：从 16 到 1024 个报文，步进为 16 个报文；

2) $\pi_0=0.01\sim0.6$；$\beta=1\sim8$；

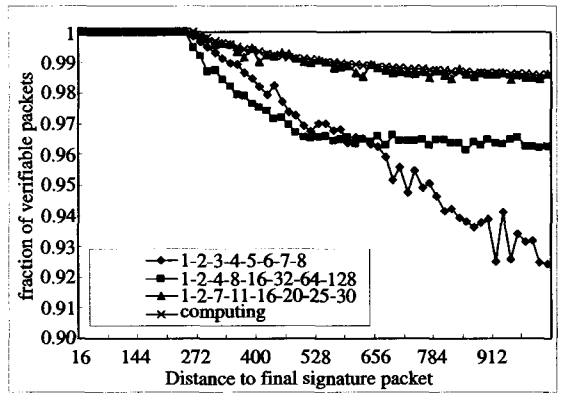
3) $k=1,2,3$；$c=2,4,8$。

4)签名报文携带的 Hash 个数(n_s)：Max(r_j)，即跨度组合的最大值；

5) Hash 跨度组合 $R=\{1,2,3,4,5,6,7,8\}$；$\{1,2,4,8,16,32,64,128\}$；$\{1,3,5,7,9,11,13,15\}$，…

与文献[4,7]相似，重复测试和比较很多种跨度组合，包括连续跨度组合 1-2-3-4-5-6-7-8，指数跨度组合 1-2-4-8-16-32-64-128 和其它很多组合。最后我们发现了一个最佳跨度组合 1-2-7-11-16-20-25-30。图 3 给出了部分组合的测试结果和根据公式(1)得出的计算结果。具体参数：$\pi_0=0.6$，$\beta=8$，$c=8$。

为了方便比较和选择不同组合的优劣，测试固定 $n_s=128$。连续跨度组合的认证概率最低；指数跨度组合的认证概率在 $n<650$ 时下降最快，但是大于 650 时下降缓慢，几乎与 n 无关；组合 1-2-7-11-16-20-25-30 的认证概率最高，在 0.99 附近，曲线下降平缓。计算结果(case 4)与该跨度组合仿真结果较为吻合，这说明该跨度组合已经逼近最优。



case 1：1-2-3-4-5-6-7-8；case 2：1-2-4-8-16-32-64-128；case 3：1-2-7-11-16-20-25-30；case 4：计算结果

图 3 跨度参数组合比较

4 性能比较

研究者们提出了很多组播认证方案，几种典型的、较为高效的方案如认证树[2]、多链方案[5-7]和冗余纠删码[8,9]。这几种方案都是采用签名开销分摊机制，并且可以容忍报文丢失认证。

认证树方案的认证概率甚为完美，但是通讯开销最高。当 n 较大时，通讯开销难以忍受。虽然可以引入概率认证方案，还是难以有效降低通信开销。SAIDA [8]是 Hash 表与纠删码结合方案的典型代表，有效地降低了通讯开销，同时解决了签名信息的有效发放问题，但是这种方案引入的延迟较大，即使在没有丢包的情况下，同时引入的发送延迟和接受延迟都很大。另外，恶意报文插入攻击抵抗力较弱，将大大增加接受端验证延迟。文献[9]改进了对恶意报文的抵抗能力，但引入了更多的延迟。

高效多链流签名(EMSS)方案[5]是经典的多链认证方案，总延迟为 n。没有发送端延迟，但是接受端延迟很大。可以将 EMSS 认证链反序，将接受端延迟转为发送端延迟，此时接受端延迟为零。HTC 方案增加了少量的发送端延迟(一个群)，但是比 EMSS 认证效率要高得多。与 EMSS 系统相同，HTC 也可以反向应用，此时没有接收端延迟，发送端延迟与 EMSS 相同。

EMSS 和 HTC 的认证概率比较如图 4 所示，图中显示了随着 n 的增加认证概率的变化。仿真测试采用与文献[5]相同的环境，具体参数如下。

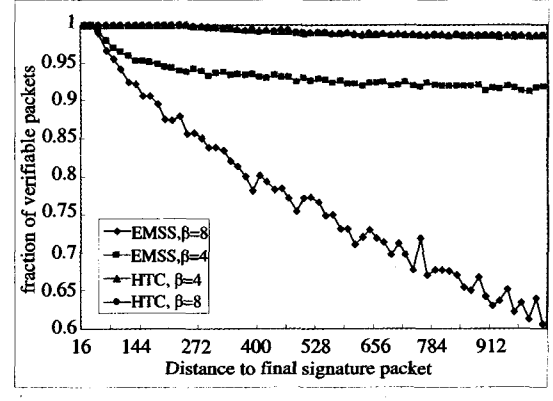


图 4 EMSS 与 HTC 验证概率的比较

维普资讯 http://www.cqvip.com

1)共同参数:Max{n}=1024,突发丢包平均长度 $\beta=4$,8,平均丢失率 $\pi_0=0.6$;

2)EMSS:签名报文携带 Hash(39),数据报文携带 Hash(6),最优组合 5-11-17-24-26-39;

3)HTC:签名报文携带 Hash(30),数据报文携带 Hash(4),最优组合 1-2-7-11-16-20-25-30。

随着 n 的增加,EMSS 在 $\beta=4$ 时,认证概率还可以接受,与文献[5]仿真结果相同,在 $\beta=8$ 时,认证概率急剧下降而无法忍受。HTC 在 $\beta=4,8$ 时认证概率几乎相同,且远高于 EMSS,并且下降并不明显,说明对报文突发性丢失的抵抗能力远高于 EMSS。此时 HTC 通讯开销远小于 EMSS:每数据报文携带 Hash 数——6(EMSS):4(HTC);签名报文携带 Hash 数——39(EMSS):30(HTC)。

结束语 本文讨论了安全组播源认证问题。大量的模拟测试证明,在突发丢包网络环境下,HTC 是一种有效的源认证方案。比较 EMSS 多链认证方案,HTC 的认证效率有了显著的提高。

参 考 文 献

[1] Rohatgi R. A compact and fast hybrid signature scheme for multicast packet authentication // Proceedings of the 6th ACM Conference on Computer and Communications Security. Singapore, November 1999:93-100

[2] Wong C K, Lam S S. Digital signatures for flows and multicasts. IEEE/ACM Transactions on Networking, 1999, 7: 502-513

[3] Paxson V. End-to-End Internet Packet Dynamics. IEEE/ACM Transactions on Networking, 1999, 7:277-292

[4] Pannetrat A, Molva R. Efficient multicast packet authentication // Proceeding of 10th Annual Network and Distributed System Security Symposium. February 2003-Symposium, 2003

[5] Perrig R C A, Song D, Tygar D. Effcient and secure source authentication for multicast // Proceedings Network and Distributed System Security Symposium (NDSS '01). San Diego, CA, Feb. 2001

[6] Golle P, Modadugu N. Authenticating Streamed Data in the Presence of Random Packet Loss // NDSS'01: The Network and Distributed System Security Symposium. 2001

[7] Miner S, Staddon J. Graph-based authentication of digital s-treams // IEEE Symposium on Security and Privacy. May 2001: 232-246

[8] Min J P, Chong E K P, Siegel H J. Efficient multicast packet authentication using signature amortization // Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy. 2002:227-240

[9] Lysyanskaya A, Tamassia R, Triandopoulos N. Multicast authentication in fully adversarial networks // IEEE Symposium on Security and Privacy. 2004

(上接第 98 页)

[3] Pendakaris D, Shi S. ALMI: an application level multicast infrastructure [A] // Anderson T, ed. The 3rd USENIX Symposium on Internet Technologies and Systems [C]. San Francisco, CA, USA: USENIX Association, 2001:49-60

[4] Chawathe Y. Scattercast: an architecture for internet broadcast distribution as an infrastructure service [D]. USA: University of California, Berkeley, 2000

[5] Francis P. Yoid: extending the multicast internet architecture [EB/OL]. http://www.aciri.org/yoid, 1999

[6] Zhang Bei-chuan, Jamin S, Zhang Li-xia. Host multicast: a framework for delivering multicast to end users [A]. Kermani P, ed. IEEE INFOCOM 2002 [C]. New York, NY, USA: IEEE Press, 2002: 1366-1375

[7] Banerjee S, Bhattacharjee B, Kommareddy C. Scalable application layer multicast [J]. ACM SIGCOMM Computer Communication Review, 2002, 32(4):205-217

[8] Zhuang S Q, Zhao B Y, Joseph A D. Bayeux: an architecture for scalable and fault-tolerant wide-area data dissemination [A] // Nieh J, Schulzrinne H, eds. The Eleventh International Workshop on Network and Operating System Support for Digital Audio and Video [C]. New York, USA: ACM Press, 2001: 11-20

[9] Castro M, Druschel P, Kermarrec A M, et al. SCRIBE: A large-scale and decentralized application-level multicast infrastructure [J]. IEEE Journal of Selected Areas in Communications, 2002, 20(8):1489-1499

[10] Ratnasamy S, Handley M, Karp R, et al. Application-level multicast using content-addressable networks [A] // Crowcroft J, Hofmann M, eds. Networked Group Communication, Third International COST264 Workshop, NGC 2001 [C]. London, UK: Springer, 2001: 14-29

[11] Wang J, Yurcik W. A survey and comparison of multi-ring techniques for scalable battlespace group communications [C] // Proceedings of SPIE, 2005-ncassr.org, Page 1

[12] Wang J, Yurcik W. Multiring techniques for scalable battlespace group communications [J]. Communications Magazine, IEEE, 2005, 43(11):124-133

[13] Junginger M, Lee Y. The multi-ring topology-high-performance group communication in peer-to-peer networks [C] // 2nd International Conference on Peer-to-Peer Computing (P2P'02). Washington, DC, USA, 2002:49-56

[14] Sobeih A, Yurcik W, HouJ C. VRing: a ring-based application-layer multicast protocol [R]. Technical Report. UIUCDCS-R-2004-2468. University of Illinois at UrbanaChampaign, 2004

[15] Sobeih A, Yurcik W, Hou J C. VRing: a case for building application-layer multicast rings [C] // Proceedings of the IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'04). Washington, DC, USA, 2004: 437-446

[16] Wang J, Yurcik W. A multi-ring framework for survivable and secure group communications [C] // Command and Control Research and Technology Symposium (CCRTS). San Diego, CA, 2004

[17] Sobeih A, Wang Jun, Yurcik W. Performance evaluation and comparison of tree and ring Application-layer multicast overlay networks [C] // Proc. of ICENCO'04. 2004

[18] Sobeih A, Yurcik W. A survey of ring-building network protocols suitable for command and control group communications [C] // Proceedings of the SPIE. 2005:873-884

[19] 王晓东. 算法设计与分析[M]. 北京: 清华大学出版社, 2003: 329-331

[20] Franklin R. On an improved algorithm for decentralized extrema finding in circular configurations of processors [J]. Communications of the ACM, 1982, 25(5): 336-337