

基于 ACK 和 NAK 的可靠组播传输协议的性能分析和比较

王 斌,刘增基,李红滨,张 冰

(西安电子科技大学综合业务网国家重点实验室,陕西西安 710071)

摘 要: 本文在目前各种可靠组播传输协议的基础上,归纳构建了两类基本的可靠组播传输协议—基于确认(ACK)的和基于否定确认(NAK)的可靠组播传输协议.在典型的组播网络系统模型基础上,对这两类协议以及一种主要的改进协议在占用网络带宽和可扩展性方面的性能特点进行了定性的分析和比较.性能分析结果表明,基于否定确认且采用否定确认压缩策略(NAK Compression)的可靠组播传输协议成功传输组播数据包所需要的带宽较小,具备良好的可扩展性.

关键词: 可靠组播传输协议; 确认/否定确认; 否定确认压缩策略

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2001) 10-1314-05

Performance Analysis and Comparison of Reliable Multicast Transport Protocol Based on ACK and NAK

WANG Bin, LIU Zeng-ji, LI Hong-bin, ZHANG Bing

(National Key Lab On Integrated Services Network, Xidian University, Xi'an, Shanxi 710071, China)

Abstract: Two types of basic reliable multicast transport protocol (ACK based and NAK based) are presented in the paper. Qualitative analysis of these two basic protocols and a modified protocol is given to exhibit their characters in scalability and bandwidth usage on a typical multicast network model. The analysis result shows that the reliable multicast transport protocol based on NAK with NAK Compression scheme consumes less bandwidth and exhibits perfect scalability.

Key words: reliable multicast transport protocol; ACK/NAK; NAK compression scheme

1 引言

迄今为止,有关可靠组播传输协议(Reliable Multicast Transport Protocol, RMTTP)的研究已经有了很多成果^[1~8].评价一个可靠组播传输协议性能的关键指标是它的可扩展性,其实质就是成功传输组播数据包时,发送者和接收者所需要的网络带宽.组播应用中有一条重要原理:不存在适用于所有应用(one size fits all)的可靠组播传输,这就是说,不同的上层应用会要求不同的可靠组播传输特性,从这个意义上讲,可靠组播协议必然是具备多样性的,相互之间很难真正进行性能的定性比较.也正因为如此,上述这些研究所提出的算法或协议的性能研究都是利用仿真方法实现的,很少有人对它们在网络带宽和可扩展性方面进行定性的分析和比较.

虽然上述各种可靠组播协议各有特点,但它们也具备一些共性.认真研究现有的各种可靠组播传输协议可以看到,它们都归结于两类基本的协议模型,本文中称之为基于确认(ACK)的和基于否定确认(NAK)的可靠组播协议.目前绝大

部分可靠组播协议都是在这两种基本协议模式的基础上变化和改进得到的.

本文在典型的组播网络系统模型基础上,对这两类协议以及一种主要的改进协议,在占用网络带宽和可扩展性方面的性能特点进行了定性的分析和比较.性能分析结果表明,基于否定确认且采用否定确认压缩(NAK Compression)策略的可靠组播传输协议成功传输组播数据包所需要的带宽较小,具备良好的可扩展性.

本文首先介绍这两种基本协议模型和一种主要的改进模型以及系统模型;然后对这些协议模型进行了详细的数学分析,并比较了它们各自的性能特点;最后得出了结论.

2 协议和系统模型

2.1 基于 ACK 的可靠组播协议模型

基于 ACK 的可靠组播协议主要由发送者来保证组播数据包的可靠传输,在本文中将此协议称为协议模型 A,它具有以下基本特点:

⑧ 发送者需要为每一接收者保留各自关于数据包接收情况的相关状态信息,它根据每次收到的来自于不同接收者的 ACK 设置相应的状态信息。

⑨ 发送者采用组播方式发送和重传每一数据包,并同时设置相应的重发定时器。

⑩ 接收者首次正确收到数据包后,利用单播 (Unicast) 方式向发送者发送 ACK 确认数据包,对于此后重复收到的数据包其不再发送 ACK 进行确认。

⑪ 若在发送者中对任一数据包的定时器超时之前,发送者收到来自所有接收者对此数据包的 ACK,则其认为此数据包已经成功的传送到所有接收者,此时取消对这一数据包的定时。

⑫ 若发送者中对应某一数据包的定时器超时,这意味着这一数据包没有正确的传送到所有接收者,则其重发此数据包并重新设置定时器。

2.2 基于 NAK 的可靠组播协议模型

基于 NAK 的可靠组播传输协议将保证组播数据包的可可靠传输的重任放在接收者上,在本文中此类协议称为协议模型 B ,它具备以下基本特点:

⑬ 发送者不需要了解每一接收者的数据包接收情况及保存相关状态信息。

⑭ 发送者采用组播方式发送和重传每一数据包,不需要设置相应定时器。

⑮ 若接收者正确收到数据包,则其不向发送者发送任何确认指示。

⑯ 接收者通过检测数据包空隙 (Packet Gap) 或通过比较各个接收者周期发送的状态信息 (本文中不考虑此状态信息所占用的带宽) 检测数据包的丢失,若接收者检测到数据包的丢失,则其利用单播方式向发送者发送 NAK 数据包,要求发送者重发相应的数据包,并同时设置相应的定时器。

⑰ 发送者收到接收者的 NAK 后,依据其中的信息决定是否重发相应的数据包,设此接收者和发送者之间的往返行程迟延为 RTT,若发送者在当前时间之前的 RTT 时间之内已经重发了相应的数据包,则发送者不再进行重发,否则其重发此 NAK 所请求的组播数据包。

⑱ 若接收者在对应某一数据包的定时器超时前成功收到此数据包,则其取消相应的定时器。

⑲ 若接收者中对应某一数据包的定时器超时,则这意味着接收者要求发送者重发的数据包还是没有成功收到,则按照检测到数据包丢失时的情况处理。

此外,现在很多基于 NAK 的可靠组播协议都采用了 NAK Compression 策略,这类协议的代是 Sally Floyd 和 Van Jacobson 的 SRM^[9],本文将其单独列为一类协议模型,称为协议模型 C ,它与协议模型 B 有一些差别:

⑳ 接收者检测到数据包的丢失后,首先按照某种退避算法等待一段时间,若这段时间内它没有收到其它接收者发送的关于同一数据包的 NAK,则其利用组播方式向所有接收者和发送者发送 NAK。

㉑ 如果接收者检测到数据包丢失并进入退避等待阶段时

收到其他接收者发送的关于同一数据包的 NAK,则其退出退避等待阶段,设置相应的定时器,就好像它已发送了 NAK 一样 (NAK Compression 策略)。

2.3 系统模型

本文针对一点对多点类型的组播应用构建了组播网络系统模型,如图 1 所示,在此模型中存在一个发送者和 R 个接收者。发送者发送的组播数据包首先到达与骨干网相连的骨干网路由器 (Backbone Router),然后通过骨干网传输到达与接收者相连的骨干网路由器,最后到达每一接收者。

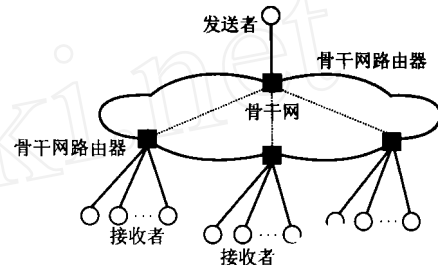


图 1 组播网络系统模型

Yajnik 等指出^[10],骨干网中组播数据包很少发生丢失,若从发送者到骨干网之间的链路丢包率很小的话,在接收者所经历的数据包丢失是相互独立的。本文系统模型中采用此结论,假定在数据包由发送者向接收者传输过程中,同一接收者上以及接收者之间所经历的数据包丢失事件均是相互独立的,其丢包率为 p 。另外,本文还假定所有 ACK 及 NAK 均不丢失。

3 协议模型的数学分析

本文中,由发送者发送的任意一个组播数据包被所有接收者正确接收称为组播数据包的成功传输。由于影响网络性能的瓶颈主要是网络带宽问题而非用户终端处理能力,因而本文研究的重点是发送者成功传输任意一个数据包时,发送者和接收者所需的平均带宽以及它们之间的关系。表 1 给出了三种协议模型中通用符号的意义解释。

表 1 三种协议模型中通用符号的意义

$B_S B_R$	成功传输任意一个组播数据包时发送者/任意一个接收者所需要的总带宽 $\{A, B, C\}$ 。
R	接收者的个数。
M	成功传输任意一个组播数据包时发送者所需发送 (重传) 数据包的次数。
$B_r(m)$	发送者/接收者每次发送 (重传)/接收数据包时所需带宽,假设它们之间相互独立,且均服从一般分布。
$B_a(n) B_n(n)$	发送者/接收者每次接收/发送 ACK/NAK 所需带宽,假设它们之间相互独立,且均服从一般分布。

3.1 基于 ACK 的可靠组播协议带宽性能分析

在协议模型 A 中,成功传输任意一个组播数据包时,发送者所需要的带宽 B_S^A 为:

$$B_S^A = \sum_{m=1}^M B_r(m) + \sum_{n=1}^{L^A} B_a(n)$$

由于本文研究的重点是发送者和接收者成功传输任意一个组播数据包所需要的平均带宽,

$$E[B_S^A] = E[M]E[B_t] + E[L^A]E[B_a]$$

其中 L^A 为发送者收到的 ACK 总数, $E[L^A] = L^A = R$

$$E[B_S^A] = E[M]E[B_t] + RE[B_a] \quad (1)$$

设 M_r 是发送者确保接收者 r 正确收到任意一个组播数据包所需发送数据包的总数.

$$P[M_r = m] = p^{m-1}(1-p), m = 1, 2, 3 \dots$$

$$P[M_r \leq m] = \sum_{k=1}^m P[M_r = k] = \sum_{k=1}^m p^{k-1}(1-p) = 1 - p^m$$

$$\text{且 } E[M_r] = \sum_{m=1}^{\infty} m \cdot P[M_r = m] = \sum_{m=1}^{\infty} P[M_r \geq m]$$

$$= \sum_{m=1}^{\infty} (1 - P[M_r \leq m - 1]) = 1/(1-p)$$

由于所有接收者之间丢包事件是相互独立的, 所以有

$$P[M \leq m] = \prod_{r=1}^R P[M_r \leq m] = (1 - p^m)^R$$

$$\begin{aligned} E[M] &= \sum_{m=1}^{\infty} m P[M = m] = \sum_{m=1}^{\infty} (1 - P[M \leq m - 1]) \\ &= \sum_{m=1}^{\infty} [1 - (1 - p^{m-1})^R] \end{aligned} \quad (2)$$

在协议模型 A 中, 成功传输任意一个组播数据包时, 任意一个接收者 r 所需要的带宽 B_R^A 为:

$$B_R^A = \sum_{m=1}^{M_r} B_t(m) + B_a(n)$$

其中 M_r 为成功传输任意一个组播数据包时, 接收者 r 总共收到的数据包数, 则接收者 r 所需要平均带宽

$$\begin{aligned} E[B_R^A] &= E[M_r]E[B_t] + E[B_a] \\ E[M_r] &= E[M](1-p) \\ E[B_R^A] &= E[M](1-p)E[B_t] + E[B_a] \end{aligned} \quad (3)$$

3.2 基于 NAK 的可靠组播协议带宽性能分析

在协议模型 B 中, 成功传输任意一个组播数据包时, 发送者所需要的带宽 B_S^B 为:

$$B_S^B = \sum_{m=1}^M B_t(m) + \sum_{n=1}^{L^B} B_n(n)$$

其中 L^B 为发送者收到的 NAK 总数. 发送者需要的平均带宽

$$E[B_S^B] = E[M]E[B_t] + E[L^B]E[B_n]$$

设 L_r^B 为接收者 r 向发送者发送的 NAK 总数 $E[L_r^B] = E[M_r] - 1 = p/(1-p)$ 且各个接收者相互独立 $E[L^B] = Rp/(1-p)$ 将其代入上式有

$$E[B_S^B] = E[M]E[B_t] + Rp/(1-p)E[B_n] \quad (4)$$

在协议模型 B 中, 成功传输任意一个组播数据包时, 任意一个接收者 r 所占用的带宽 B_R^B 为:

$$B_R^B = \sum_{m=1}^{M_r} B_t(m) + \sum_{n=1}^{L_r^B} B_n(n)$$

发送者占用的平均带宽为:

$$E[B_R^B] = E[M_r]E[B_t] + E[L_r^B]E[B_n]$$

$E[M_r] = E[M](1-p)$, $E[L_r^B] = p/(1-p)$ 将其代入上式有

$$E[B_R^B] = E[M](1-p)E[B_t] + p/(1-p)E[B_n] \quad (5)$$

在协议模型 C 中, 成功传输任意一个组播数据包时, 发

送者所占用的带宽 B_S^C 为:

$$B_S^C = \sum_{m=1}^M B_t(m) + \sum_{n=1}^{L^C} B_n(n)$$

其中 L^C 为发送者收到的 NAK 总数, 发送者占用的平均带宽

$$E[B_S^C] = E[M]E[B_t] + E[L^C]E[B_n]$$

假设 NAK Compression 策略工作情况理想, 此时有 $E[L^C] = E[M] - 1$ 代入上式中有:

$$E[B_S^C] = E[M]E[B_t] + (E[M] - 1)E[B_n] \quad (6)$$

在协议模型 C 中, 成功传输任意一个组播数据包时, 任意一个接收者 r 所需要的带宽 B_R^C 为:

$$B_R^C = \sum_{m=1}^{M_r} B_t(m) + \sum_{n=1}^{L_r^C} B_n(n) + \sum_{i=1}^{L_r^C} B_n(i)$$

其中 L_r^C 为接收者 r 自己发送 NAK 的数目, L_r^C 为接收者 r 从其它接收者收到 NAK 的数目, $B_n(i)$ 为接收者 r 每次接收从其它接收者来到的 NAK 时占用的带宽. 接收者 r 需要的平均带宽为:

$$E[B_R^C] = E[M_r]E[B_t] + E[L_r^C]E[B_n] + E[L_r^C]E[B_n]$$

$$E[L_r^C] = (E[M] - 1)/R, E[L_r^C] = (E[M] - 1)(R - 1)/R$$

且 $E[M_r] = E[M](1-p)$ 将它们代入上式中有:

$$\begin{aligned} E[B_R^C] &= E[M](1-p)E[B_t] + (E[M] - 1)/RE[B_n] \\ &\quad + (E[M] - 1)(R - 1)/RE[B_n] \end{aligned} \quad (7)$$

将式(2)代入式(1)及式(3~7)就可以得到成功传输任意一个组播数据包时三种协议模型中发送者和接收者各自所需要的平均带宽.

比较式(1)、(4)和(6), (3)、(5)和式(7)可以明显的看出, 在假设 ACK/NAK 不发生丢失的前提下, A、B、C 三种协议模型中成功传输一包数据包时发送者/接收者上发送/重发数据包占用的带宽量是一样的, 它们的差别在于反馈信息(ACK/NAK)所占用的带宽, 实际上就是反映在 $R, Rp/(1-p), E$

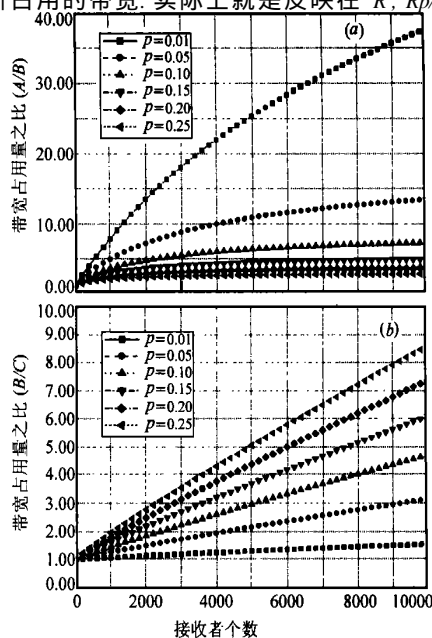


图2 协议模型 A, B, C 所需平均总带宽之间的关系

$[M] - 1$ 三者之间的关系上. 当 $p < 0.5$ 时(这也符合实际网络情况), 有 $R > Rp/(1-p) > E[M] - 1$, 也就是说协议模型 C 的带宽占用量最小, B 次之, A 最大.

4 协议模型的性能比较

上一节中给出了各种协议模型的数学分析, 这一节将就这些协议模型成功传输任意一个数据包时发送者和接收者所需平均带宽及其关系进行研究比较. 由于可扩展性是评价一个可靠组播传输协议的重要指标, 发送者和接收者所需平均带宽随组播规模(接收者个数)变化时的特性正是其可扩展性的体现, 因而本文也比较了它们在可扩展性上的差异. 依据文[10]中在 MBONE 中观测得到的数据, 本文中取 $E[B_a] = E[B_n] = E[B_r] = 32$, $E[B_t] = 1024$, $p \in [0.01, 0.25]$, $R \in [1, 10000]$.

图 2 显示了协议模型 A 、 B 、 C 在不同组播规模和数据包丢失率 p 情况下需要的平均总带宽之间的关系. 从图 2(a) 中可以看出, 当接收者个数保持不变时, 随着 p 的增加, A/B 越来越小. 这是因为在传输链路条件恶化, 丢包率增大的情况下, 所有接收者上的丢失的数据包增多, 因此协议模型 A 中 ACK 需要的带宽相应减小而协议模型 B 中 NAK 需要带宽则相应增大, 所以导致了 A/B 减小. 当 p 保持不变时, 随着组播规模的增大, A/B 越来越大, 这是由于协议模型 A 中 ACK 的大量增加造成的. 反映了协议模型 A 在可扩展性上存在的缺

陷. 因而, 从总体上讲, 协议模型 B 的性能要优于 A . 图 2(b) 中显示了 B 、 C 之间在需要的平均总带宽之间的关系, 有一点值得特别注意的是, 随着 p 的增大, 协议模型 B 中 NAK 所占用的带宽相应的增大. 而协议模型 C 由于采用了 NAK Compression 策略因而它的 NAK 所占用的带宽的只有少量的增大, 因而导致了 B/C 的增大. 从可扩展性方面看, 也可以看出协议模型 C 要优于 B .

图 3 显示了协议模型 A 、 B 、 C 在不同组播规模和数据包丢失率 p 情况下, 发送者和接收者需要的平均带宽的比值(发送者/接收者). 在图 3(a) 中, 当组播规模保持不变时, 随着 p 逐渐变大, 发送者与接收者带宽占用量的比值逐渐变小, 这是由此时发送者上反馈的 ACK 数量减少, 同时接收者上的带宽占用量基本不变而造成的. 当 p 保持不变时, 随着组播规模的增大, 发送者会收到越来越多的反馈 ACK, 此时在发送者附近就很容易造成拥塞从而会影响网络和组播应用的性能. 如图 3(b), 在协议模型 B 中由于反馈信息 NAK 所需要的平均带宽要远小于 A 中 ACK 所需的平均带宽, 因而这一值就远远小于协议模型 A 在相同网络条件下的值. 从而大大降低了在发送者附近造成拥塞的可能性. 如图 3(c), 协议模型 C 由于采用了 NAK Compression, 从而进一步的减少了 NAK 所需要的平均带宽, 此时此值很小, 而且基本上不随组播规模的增大而变化, 体现了其具备的良好可扩展性.

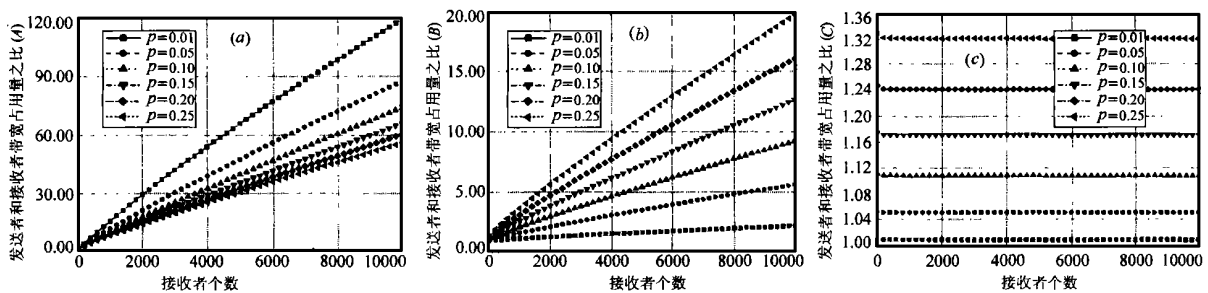


图 3 协议模型 A 、 B 、 C 中发送者和接收者带宽量之间的关系

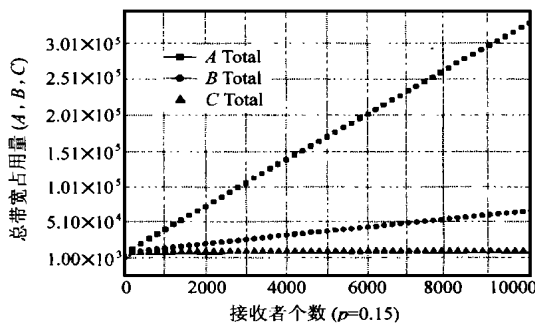


图 4 三种协议模型所需平均总带宽的比较

图 4 显示了 $p=0.15$ 时协议模型 A 、 B 、 C 所需要的平均总带宽. 从图中可以明显的看出, 随着组播规模的增大, 协议模型 A 所需要的带宽变得非常大, 因而它的可扩展性是非常差的. 总体上讲, 采用 NAK Compression 的协议模型 C 的可扩展性是三者中最好的.

5 结论

本文在典型的组播网络系统模型基础上, 定性的分析了两类基本的可靠组播传输协议和一种主要的改进协议. 这些协议模型是当前所有可靠组播传输协议的基础. 通过数学分析和性能比较可以看出, 基于 NAK 且采用 NAK Compression 策略的可靠组播传输协议成功传输组播数据包时所需要的带宽较小, 当组播规模增大和丢包率变大时, 仍能保持很低的带宽需求, 显示了其具备良好的可扩展性, 可以作为今后可靠组播传输协议研究方向的基础.

参考文献:

- [1] P Bhagwat, P Misra, S Tripathi. Effect of topology on performance of reliable multicast communication [A]. Proc. IEEE Infocom 94 [C], Toronto, June 1994:602 - 609.
- [2] A Erramilli, R P Singh. A reliable and efficient multicast protocols for broadband broadcast networks [A]. Proc. ACM Sigcomm88 [C], August 1988:343 - 352.

- [3] H W Holbrook ,S K Singhal ,D R Cheriton. Log-Based receiver reliable multicast for distributed interactive simulation [A]. Proc. of ACM Sigcomm [C] ,August 1995 :328 - 341.
- [4] S K Kasera J Kurose ,D Towsley. Scalable reliable multicast using multiple multicast groups [A]. Proc. of ACM Sigmetrics '97 [C] ,Seattle ,WA ,1997 :64 - 74.
- [5] B N Levine ,D B Lavo ,J J Garcia-Lunar-Aceves. The case for reliable concurrent multicasting using shared ack trees [C]. Proc. of Multimedia ,1996 :365 - 376.
- [6] J C Lin ,S Paul. RMTP: A reliable multicast transport protocol [A]. Proc. IEEE Infocom [C] ,March 1996 :1414 - 1425.
- [7] J Nonnenmacher ,E Biersack ,J Towsle. Parity-based loss recovery for reliable multicast transmission [R]. Technical Report ,TR-97-17 ,Department of Computer Science ,University of Massachusetts ,Mar 1997.
- [8] Christos Papadopoulos ,George Varghese. An error control scheme for large scale multicast applications [A]. Proc. of the IEEE Infocom [C] ,March 1998 :1188 - 1196.
- [9] S Floyd ,V Jacobson ,S McCanne ,C G Liu ,L Zhang. A reliable multicast framework for light-weight sessions and application-level framing [A]. Proc. ACM Sigcomm95 [C] ,Boston MA ,Aug. 1995 :342 - 356.

- [10] M Yajnik ,J Kurose ,D Towsley. Packet loss correlation in the MBone multicast network [A]. IEEE Global Internet '96 [C] ,London ,England ,November 20 - 21 ,1996 :217 - 225.

作者简介:

王 斌 男. 1973 年 6 月生于陕西. 分别于 1995 年和 1998 年在西安电子科技大学通信与信息系统专业获学士、硕士学位, 研究方向为综合业务数字网与 IP 网络. 从 1998 年至今, 于西安电子科技大学 ISN 国家重点实验室继续攻读博士学位, 研究方向包括主动式网络, 可靠组播传输, QoS 以及 ISDN.



刘增基 男. 生于 1937 年 12 月, 现为西安电子科技大学教授、博士生导师、综合业务网国家重点实验室主任、中国通信学会会士. 当前主要从事宽带通信网络技术的研究.

中国电子学会第八届青年学术年会 征 文 通 知

中国电子学会第八届青年学术年会(简称:CIE-YC'2002)将于 2002 年 8 月在安徽合肥召开. 这是一次团结广大青年科技工作者, 促进电子信息及其相关学科青年学者学术交流的盛会, 届时还将邀请国内著名专家学者作综述或专题报告, 大会将在参会宣读论文中评选出优秀论文, 并推荐到《模式识别与人工智能》杂志上发表.

征文范围(但不限于这些领域):

- | | |
|-------------------|----------------|
| A: 信息技术发展与展望 | B: 微电子技术及其应用 |
| C: 纳米与量子信息技术 | D: 系统工程与应用 |
| E: 机器视觉与计算机图形学 | F: 雷达、通信与电子战技术 |
| G: 数字信号处理 | H: 卫星、广播电视技术 |
| I: 农业信息技术与生长模型 | J: 多媒体与虚拟现实 |
| K: 计算机技术与应用 | L: 器件、电路与系统 |
| M: 模式识别与图像处理 | N: 神经网络与模糊软计算 |
| O: 电子技术及应用 | P: 光电子技术与电子材料 |
| Q: 传感器、机器人与自动控制技术 | R: 电磁场与微波技术 |
| S: 数据融合、挖掘与知识发现技术 | |
| T: 人工生命与生物(基因)信息学 | U: 非线性信号处理技术 |

来稿要求:

1. 内容具体, 突出作者的创新与成果, 具有较重要的学术价值和推广应用价值, 未在国内公开发行的刊物或会议上发表或宣读过.

2. 第一作者年龄不超过 40 岁, 文末附不超过 100 字的第一作者简介.

3. 来稿时请在论文背后注明论文所属征文范围中的类别字母 1 至 2 个, 并在信封上注明征文字样.

4. 录用的论文拟由中国科技大学出版社正式出版论文集. 投稿方式为全文投稿方式, 每篇论文篇幅限为 4 页(超过者加收版面费). 寄送论文请用 WORD 排版, 激光打印, 投稿时请寄送两份, 并要求随文附寄软盘. 论文第一页正文前应写: 论文题目、作者姓名、作者单位、邮政编码; E-mail 地址及摘要. 论文排版版式样张附后.

重要时间:

论文截止时间: 2002 年 4 月 15 日

录用通知: 2002 年 5 月 15 日前发出

会议时间: 2002 年 8 月

投稿地址:

230031 安徽省合肥市科学岛 1130 信箱 张海梅 黄德双
电 话: (0551) 5591108 E-mail: 2002CIEYC@mail.iim.ac.cn
传 真: (0551) 5592420 网 页: http://www.iim.ac.cn/

中国电子学会青年工作委员会

中科院合肥智能机械研究所

二 一 年九月六日