

随机弱实时系统作业时限延误率上界估计

彭良智,戴金海,单 懿,桂先洲

(国防科技大学航天与材料工程学院,长沙 410073)

摘 要: 本文提出了关于随机弱实时系统的统计处理器需求分析 (SPDA) 方法,该方法可用于估计可剥夺静态优先级调度策略下作业满足时限约束的概率的下界.给出了 SPDA 的数值解法.对实例系统的仿真结果表明,SPDA 分析的误差小于 6%,计算速度快,可以为弱实时应用的设计提供重要的参考依据.

关键词: 随机弱实时系统; SPDA 分析; 下界估计

中图分类号: TP391.9 **文献标识码:** A **文章编号:** 0372-2112 (2001) 02-0186-04

Upper Bound Estimation of Job Deadline Missing Rate for Stochastic Soft Real-Time Systems

PENG Liang-zhi, DAI Jin-hai, SHAN Yi, GUI Xian-zhou

(College of Astronautics and Material Engineering, National University of Defense Technology, ChangSha 410073, China)

Abstract: This paper proposes statistical processor demand analysis (SPDA) scheme for stochastic soft real-time system. The scheme can be used to estimate the lower bound on the probability that a job meets its deadline under fixed-priority preemptive schedule. Numerical solution to SPDA is also presented. The simulation results of an example system have shown that SPDA has a high computing speed and an error below 6%. It can provide important reference for the design of soft real-time applications.

Key words: stochastic soft real-time systems; SPDA analysis; lower bound estimation

1 引言

实时系统是能够在确定的时间内执行计算或处理事务并对外部的异步事件做出响应的计算机系统.它不但要保证任务计算结果的逻辑正确性,而且要在规定的时间内完成.这个规定的时间界限称为任务时限 (Deadline).根据任务时限的不同类型,实时系统可分为强实时、准实时和弱实时系统^[1].我们称作业对处理器资源的需求随机的弱实时系统为随机弱实时系统.这类系统在远程多媒体通信、信号处理、工业控制等领域广泛存在.以远程 VBR 动态视频点对点传输为例,让目标机底层运行实时操作系统,处理数据接收、MPEG 视频流解码、播放等实时任务,其中对解码这一周期任务而言,1 帧的解码时间明显长于 P 帧或 B 帧.假设将延误时限的帧丢弃.系统不要求每一项作业都满足时限要求,只要保证作业的时限延误率低于某一阈值 (即帧丢失率低于某一水平),其总体执行性能就是可以接受的.

对实时系统的可调度性分析主要有确定性分析和统计性分析两种.前者主要针对强实时系统,后者主要针对弱实时系统,如 Tia 等人提出的性能概率保证 (Probabilistic Performance Guarantee, PPG) 分析^[2].PPG 分析假设所有任务的相对时限不超过周期,认定任务第一项作业满足时限的概率就是整个任务时限满足率的下界,这些假设在系统的平均资源利用率接

近 1.0,最大资源利用率大于 1.0 时是不成立的.本文对 Lehoczky 提出的强实时环境下的时间需求分析法 (Time Demand Analysis, TDA)^[3]和忙周期分析法 (Busy Period Analysis, BPA)^[4]加以改造并移植到对随机弱实时系统的可调度性分析上,提出了 SPDA (Statistical Processor Demand Analysis) 分析法,用于估计可剥夺静态优先级调度策略下周期性任务的作业满足时限约束的概率的下界.SPDA 分析与 PPG 分析相比,对任务参数的限制更少,适用范围更广,计算更简便,结果更准确.设计者可将分析结果与应用要求相比较,不断改进设计.它提供了一条对弱实时系统进行事前分析的更为便捷的途径.

2 问题描述

假设随机弱实时系统由 n 个周期任务 T_1, T_2, \dots, T_n 组成,每个任务由五元组 $T_i (P_i, D_i, R_i, I_i, f_i(t))$ 表征, P_i, D_i, R_i, I_i 分别表示任务的周期、相对时限、就绪时间、初始相位, $f_i(t)$ 是作业处理需求的概率密度函数 (以下简称 PDF).任务 T_i 的第 j 个作业 $J_{i,j}$ 的处理需求为 $E_{i,j}$,它是从统计分布 $f_i(t)$ 得到的随机样本.对给定的 $i, E_{i,j}, j = 1, 2, \dots$ 独立同分布,不同任务的作业执行时间之间也是相互独立的. $J_{i,j}$ 在 $r_{i,j}$ 时刻就绪,必须在绝对时限 $d_{i,j}$ 之前完成,否则称之为延误时限.

绝对时限与就绪时间之差为任务的相对时限,即 $D_i = d_{i,j} - r_{i,j}$. 记 $J_{i,j}$ 的完成时间为 $C_{i,j}$, 则作业响应时间为 $r_{i,j} = C_{i,j} - r_{i,j}$.

假定采用可剥夺的静态优先级调度策略,记 T_i 的优先级为 ϕ_i , 作业流 $J_{i,1}, J_{i,2}, \dots$ 的优先级也相应为 ϕ_i . 本文假定任务的优先级各不相同,并按优先级降低的次序排列所有任务. $E_i^- (E_i^+)$ 、 \bar{E}_i 表示任务 T_i 的所有作业中最小(最大)和平均处理需求. 系统处理机利用率总和的平均值和最大值分别为 $U_{ave} = \frac{1}{n} \sum_{i=1}^n \bar{E}_i / P_i$ 和 $U_{max} = \frac{1}{n} \sum_{i=1}^n E_i^+ / P_i$. 在单处理机(Uniprocessor)条件下,假设 $U_{max} > 1.0$, $U_{ave} < 1.0$, 由经典的实时系统调度理论可知,即使 U_{ave} 很小,平均情况下满足系统可调度性的充分条件,但由于 $U_{max} > 1.0$, 可能出现瞬间过载,无法保证系统总是可调度的,因此有必要估计出各任务的作业延误时限的概率的上界,并据此判断是否满足具体的实时应用的要求.

3 统计处理机需求分析

3.1 SPDA 的简单情形-任务间相互独立

定义 1^[7] ϕ_i -级忙周期是一个时间区间 $[a, b]$, 在该区间内, 处理机一直无空闲地处理优先级高于等于 ϕ_i 的作业. 对 $\forall \epsilon > 0$, 在时间区间 $(a - \epsilon, a)$ 和 $(b, b + \epsilon)$ 上不处理任何优先级高于等于 ϕ_i 的作业.

定义 2^[6] 随机变量的条件映射 $h: x \rightarrow y$, 记作 $y = h(x)$ (对 x 的约束条件). 若 x 对应的 PDF 为 $f_X(x)$, 则随机变量 y 对应的 PDF 为:

$$f_Y(y) = \begin{cases} f_X(y)/A, & y \in E \\ 0, & y \notin E \end{cases}$$

其中 E 是满足指定条件的 x 的集合, $A = \int_x f_X(x) dx$.

为了确定作业时限延误率的上界, 必须找出作业响应时间的最坏情形. 由经典实时调度理论可知, 任务 T_i 的所有作业的最长响应时间发生在 ϕ_i -级忙周期内, 这个忙周期起始于临界时刻, 即所有任务的初始相位相同, 它们的第一个作业同时发布并就绪. 不失一般性, 不妨把这个 ϕ_i -级忙周期的起点当作时间原点. 下面基于这种最坏情形进行讨论.

首先计算 ϕ_i -级忙周期中 T_i 发布的第 j 个作业 $J_{i,j}$ 的响应时间的概率分布. 令 $w_{i,j}(t)$ 是需在时间区间 $(r_{i,j}, t]$ 上执行的所有作业的处理时间需求, $p_{i,j}(t) = P\{w_{i,j}(t) \leq t\}$ 表示 ϕ_i -级忙周期尚未结束时处理机时间需求在 t 之前已被满足, $J_{i,j}$ 已经完成的概率. 响应时间分布 $p_{i,j}(t)$ 与 $J_{i,j}$ 发布时是否存在优先级高于等于 ϕ_i 的作业积压有关. 对优先级最高的任务而言, ϕ_i -级忙周期中第一个作业的响应时间分布等同于作业执行时间的分布, 而其后续作业的响应时间分布可以通过计算条件作业积压的分布与作业执行时间的卷积得到.

计算 $w_{i,j}(t)$, $j > 1$ 时, 需综合考虑各种因素的影响. 优先级高于等于 ϕ_i 的作业显然可以在 $(r_{i,j}, C_{i,j})$ 区间执行; 作业 $J_{i,1}, J_{i,2}, \dots, J_{i,j-1}$ 中在 $J_{i,j}$ 发布之后完成的也可在该区间执行; 还必须考虑 $(r_{i,j}, C_{i,j})$ 期间发布的优先级高于 ϕ_i 的作业. 按优先级高于 ϕ_i 的作业发布的顺序, 把 $(r_{i,j}, C_{i,j})$ 分割成一系列的子区间, 即:

$$(r_{i,j}, C_{i,j}] = (r_{i,j}, r_{k_1, l_1}] \cup (r_{k_1, l_1}, r_{k_2, l_2}] \cup \dots \cup (r_{k_{n-1}, l_{n-1}}, r_{k_n, l_n}] \quad (1)$$

其中 r_{k_i, l_i} , $i = 1, 2, \dots, n-1$ 依次递增, 对应各高优先级作业发布的时刻, $r_{k_n, l_n} = C_{i,j}$. 若 $J_{i,j}$ 在 $t \in (r_{i,j}, r_{k_1, l_1}]$ 之前完成, 对处理机总的时间需求为:

$$w_{i,j}(t) = E_{i,j} + B_{i,j-1}(r_{i,j}) \quad (2)$$

其中 $B_{i,j-1}(r_{i,j})$ 表示作业 $J_{i,j-1}$ 截至 $r_{i,j}$ 时刻尚未完成时的条件作业积压分布, 即:

$$B_{i,j-1}(r_{i,j}) = h(w_{i,j-1}(t) | t > r_{i,j}) \quad (3)$$

此时 $J_{i,j}$ 的响应时间分布为:

$$p_{i,j}(t) = P\{w_{i,j}(t) \leq t | w_{i,j-1}(r_{i,j}) > r_{i,j}\} \quad (4)$$

若 $J_{i,j}$ 在时刻 $t \in (r_{k_i, l_i}, r_{k_{i+1}, l_{i+1}}]$, $i = 1, 2, \dots, n-1$ 之前完成:

$$B_{i,j}(r_{k_i, l_i}) = h(w_{i,j}(t) | t > r_{k_i, l_i}) \quad (5)$$

相应的时间需求和响应时间分布为:

$$w_{i,j}(t) = E_{k_i, l_i} + B_{i,j}(r_{k_i, l_i}) \quad (6)$$

$$p_{i,j}(t) = P\{w_{i,j}(t) \leq t | w_{i,j-1}(r_{i,j}) > r_{i,j},$$

$$w_{i,j}(r_{k_i, l_i}) > r_{k_i, l_i}\} \cdot P\{w_{i,j}(r_{k_i, l_i}) > r_{k_i, l_i}\} \quad (7)$$

$p_{i,j}(t)$ 的表达式中乘积的后一项是对应该子区间的概率折扣因子. 综合 $w_{i,j}(t)$ 在各子区间的分布, 得到作业 $J_{i,j}$ 在其时限之前完成的概率:

$$p = \prod_{i=1}^M p_{i,j}(r_{k_i, l_i}) \quad (8)$$

其中 $M = \max\{i | r_{k_i, l_i} \leq d_{i,j}\}$. 对 ϕ_i -级忙周期中任务 T_i 发布的所有作业, 用上述方法计算它们满足时限的下限概率, 所有这些概率的最小值即为任务 T_i 的作业满足时限的概率的下界.

其次, 确定 ϕ_i -级忙周期何时结束是 SPDA 分析的关键. 由 ϕ_i -级忙周期的定义, 可得如下结论:

定理 1 随机弱实时系统周期任务集 T_1, T_2, \dots, T_n 按优先级降序排列, 每个优先级对应一个任务, 采用可剥夺的静态优先级调度算法. 如果 $\exists j > 0$, 使得 $P\{w_{i,j}(r_{i,j+1}) \leq r_{i,j+1}\} = 1.0$, 即作业 $J_{i,j}$ 在其后续作业 $J_{i,j+1}$ 发布之前已经完成, 则 ϕ_i -级忙周期在 $J_{i,j}$ 完成的时刻结束.

推论 在定理 1 的前提但存在多个任务具有相同的优先级的情况下, ϕ_i -级忙周期在 t 时刻结束, 当且仅当: t 时刻之前发布的所有优先级高于等于 ϕ_i 的作业在 t 之前已经完成的概率为 1.0.

3.2 SPDA 的复杂情形——对资源同步的处理

实际应用中的实时系统的任务之间一般非独立, 例如, 由于资源共享, 任务之间存在间接的相互制约关系. 资源同步协议的好坏直接影响到系统的可预测性和可调度性. 以下的讨论基于临界区不剥夺 (Non-preemptible Critical Section, NPCS) 协议. 根据 NPCS 协议, 如果某作业正处于临界区访问共享资源, 其他作业不能剥夺它的执行, 直到该作业退出临界区为止. A. K. Mok 在文[5]中已证明, 作业的最大阻塞时间是所有能够阻塞它的低优先级作业中最长的临界区持续时间.

定理 2 ϕ_i -级忙周期中只有第一个作业可能被阻塞.

证明: 当没有阻塞时, 只有优先级高于等于 ϕ_i 的作业能

在 ϕ_i -级忙周期内执行,因此对优先级低于 ϕ_i 的作业,除 ϕ_i -级忙周期开始的时刻外,它不可能进入临界区,因此只有 ϕ_i -级忙周期中的第一个作业可能被阻塞。

注意到作业临界区持续时间为随机变量. 令 $B_{i,j}^k$ 为作业 $J_{i,j}$ 单独执行时,它的第 k 个临界区的持续时间, b_i^{k+} 为随机变量 $B_{i,j}^k$ 的最大值, b_i^+ 为任务 T_i 的所有 k 个临界区的持续时间的最大值,即 $b_i^+ = \max_{1 \leq m \leq k} B_{i,j}^m$. 由定理 2,只需考虑阻塞时间对 ϕ_i -级忙周期中第一个作业的处理时间需求 $w_{i,1}(t)$ 的影响,即:

$$w_{i,1}(t) = w_{i,1}(t) + \max_{i < j \leq n} b_j^+ \quad (9)$$

其中 $w_{i,1}(t)$ 是任务相互独立时 $J_{i,1}$ 的处理时间需求. 式(9)考虑了共享资源时任务 T_i 被阻塞的最坏情形. 通常情况下,为提高系统的执行性能和响应速度,临界区的持续时间都很短,相互之间的差异也不大,因此由式(9)计算出来的处理时间需求不至于过分保守。

4 SPDA 分析的数值实现

当作业执行时间的概率分布太复杂或无法用解析式表达时,必须用数值方法实现 SPDA 分析. SPDA 分析牵涉到的主要运算是两个相互独立的随机变量的 PDF 的卷积. 先对两个 PDF 进行 N 点离散化,再利用快速傅里叶变换 (FFT) 计算二者的卷积,计算复杂度为 $O(N \log_2 N)$. 需注意以下问题:第一,在对两个随机变量的 PDF 进行离散化表示时,采样率和离散点数必须相等;第二,离散化后的矢量长度必须是 2 的幂次,长度不够的在后面添 0;第三,对 PDF 离散化时,离散后的矢量长度要合适的大小. 抽样点数越多,误差越小,但计算时间相应增加. 我们的实验表明,对计算复杂度和结果的精度折衷考虑,选取 $N = 1024$ 比较合适,卷积结果的绝对误差在 0.005 以内。

5 实例分析与讨论

某随机弱实时系统由两个相互独立的周期任务 T_1, T_2 组成,如表 1 所示. 作业执行时间服从均匀分布. 平均和最坏情形下资源利用率总和分别是 0.708 和 1.411,因此在单处理机条件下,可能会有一些作业延误时限. T_1 的优先级高且它的最大资源利用率小于 1.0,故其所有作业均能满足时限约束. 只需对 T_2 进行分析.

表 1 任务属性

T_i	P_i	D_i	E_i^-	\bar{E}_i	E_i^+	U_i^-	\bar{U}_i	U_i^+
T_1	300	300	1	100	199	0.0033	0.333	0.663
T_2	400	400	1	150	299	0.0025	0.375	0.748
总和						0.0058	0.708	1.411

为计算 $J_{2,1}$ 按时完成的概率,把区间 $(0, 400)$ 分割成两个子区间 $(0, 300)$ 和 $(300, 400)$. 若 $J_{2,1}$ 在 $t \in (0, 300)$ 之前完成,处理时间需求 $w_{2,1}(t)$ 为 $J_{1,1}$ 和 $J_{2,1}$ 的执行时间之和,其 PDF 为:

$$f_z(z) = \begin{cases} \frac{1}{6 \times 10^4} z, & 0 \leq z < 200 \\ \frac{1}{300}, & 200 \leq z < 300 \\ \frac{1}{6 \times 10^4} (500 - z), & 300 \leq z \leq 500 \end{cases} \quad (10)$$

$J_{2,1}$ 在 $t = 300$ 之前完成的概率为 $P\{W_{2,1}(300) \leq 300\} = 0.668$. 若 $J_{2,1}$ 在 $t \in [300, 400)$ 之前完成,时间需求 $w_{2,1}(t)$ 是 $J_{1,2}$ 的执行时间与截至 $t = 300$ 时刻 $J_{2,1}$ 尚未完成时的条件作业积压分布 $B_{2,1}(300)$ 之和, $B_{2,1}(300)$ 是对 $w_{2,1}(t), t \in (0, 300)$ 的 $(300, 498)$ 区间进行归一化处理得到的条件映射分布, $w_{2,1}(t), t > 300$ 的 PDF 为:

$$f_z(z) = \begin{cases} -\frac{1}{8 \times 10^6} (z^2 - 1000z + 2.1 \times 10^5), & 300 \leq z < 500 \\ \frac{1}{8 \times 10^6} (z^2 - 1400z + 4.9 \times 10^5), & 500 \leq z \leq 700 \end{cases} \quad (11)$$

因此当 $J_{2,1}$ 在 $t = 300$ 之前未完成而在 $t = 400$ 之前完成的条件概率为 0.209. 综上所述, $J_{2,1}$ 在其时限 $t = 400$ 之前完成的概率为 $P\{w_{2,1}(400) \leq 400\} = 0.668 + 0.209 \times (1 - 0.668) = 0.738$. 同理可分析出 $J_{2,2}$ 在其时限 $t = 800$ 之前按时完成的概率为 0.994, $J_{2,3}$ 在其时限 $t = 1200$ 之前按时完成的概率为 1.0. ϕ_2 -级忙周期在 $J_{2,3}$ 完成时结束,故 T_2 的作业时限延误率上界为 $1 - \min(0.738, 0.994, 1) = 0.262$.

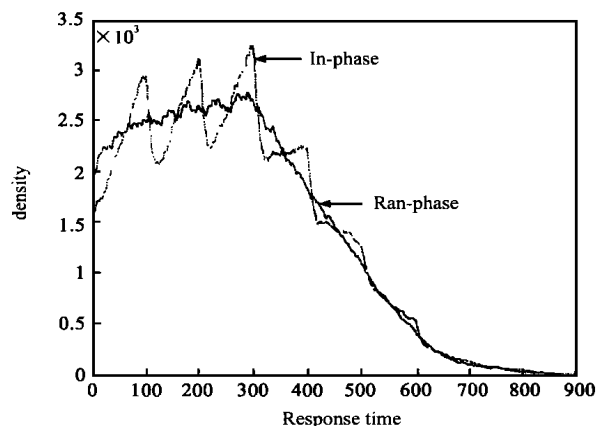


图 1 任务 T_2 作业响应时间概率密度曲线

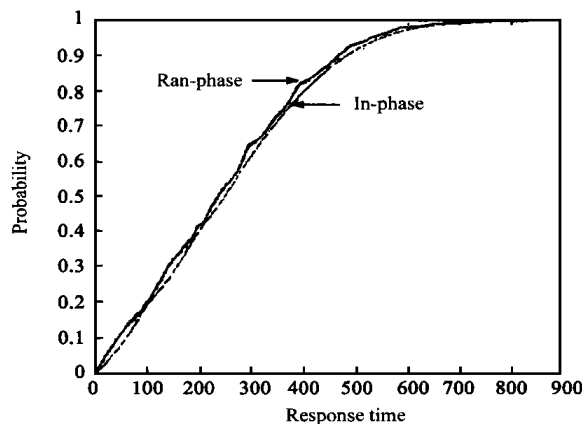


图 2 任务 T_2 作业响应时间概率分布曲线

对系统作 1000 次仿真运行,每次运行执行 3000 项 T_2 的作业和 4000 项 T_1 的作业,得到 T_2 作业响应时间的直方图,画出其概率密度和概率分布曲线,如图 1 和图 2 所示,其中初相同步 (In-phase) 情形取 T_1 和 T_2 的初始相位为 0,初相随机

(Ran-phase)情形则让 $T_i, i = 1, 2$ 的初始相位服从 $(-P_i, P_i)$ 上的均匀分布. Ir-phase 情形下 PDF 曲线在 100 的整数倍处呈现“锯齿状”,原因是 T_1 和 T_2 的作业发布时间具有固定的关系,使 T_2 的作业响应时间周期性地被延迟. Ir-phase 情形下的概率分布曲线被 Ran-phase 情形下的曲线从上面“包络”住,表明 Ir-phase 情形下总体作业响应时间更大.表 2 列出了仿真结果与 SPDA 分析、PPG 分析结果的对比,从中可看出 Ran-phase 情形下作业满足时限要求的概率高于 Ir-phase 情形. SPDA 分析得到的作业满足时限的概率与仿真结果的误差仅为 - 5.9 %.用 PPG^[6]方法分析, T_2 作业满足时限的概率为 66.7 %,与仿真结果的误差为 - 14.9 %,可见 SPDA 分析结果更准确.

表 2 仿真结果与 SPDA、PPG 分析结果的对比

T_i	SPDA 结果	误差	PPG 结果	误差	仿真实验结果	
					初相同步情形	初相随机情形
T_1	100.0	0	100.0	0	100.0 \pm 0.0	100.0 \pm 0.0
T_2	73.8	- 5.9 %	66.7	- 14.9 %	78.4 \pm 0.1	80.2 \pm 0.1

6 结束语

如果不对强实时系统分析技术进行改造就直接用于随机弱实时系统的设计,会使系统的资源利用率和总体执行性能大大降低.本文提出的单机环境下的 SPDA 分析方法是对各种类型的随机弱实时系统进行可调度性分析的基础,它估计了随机弱实时系统在任务相互独立和存在资源同步两种情形下作业时限延误率的上界,并通过具体的实例考察了上界的紧性.仿真结果表明,SPDA 分析误差小,速度快,可以为弱实时应用的设计提供重要的参考依据,尤其适合于对那些难以做仿真实验的实时系统进行事前分析.今后还将把 SPDA 分析拓展到动态优先级调度情形、任务同步情形以及分布式环境下随机弱实时系统的可调度性分析上.

参考文献:

[1] 毛羽刚. 分布强实时系统调度技术的研究 [D]. 国防科技大学

工学博士学位论文,1999. 10.

[2] T. S. Tia ,Z. Deng ,et al. Probabilistic performance guarantee for real-time tasks with varying computation times [J]. Proceedings of IEEE Real-time Technology and Applications Symposium ,1995 ,5 : 164 - 173.

[3] J. Lehoczky ,L. Sha , Y. Ding. The rate monotonic scheduling algorithms :exact characterization and average case behaviour [J]. IEEE Real-Time Systems Symposium ,1989 ,12 :166 - 171.

[4] J. Lehoczky. Fixed priority scheduling of periodic task sets with arbitrary deadlines [J]. IEEE Real-Time Systems Symposium ,1990 ,12 :201 - 209.

[5] A. K. Mok. Fundamental design problems of distributed systems for the hard real-time environment [D]. Ph. D Thesis ,MIT ,1983.

[6] Peng Liangzhi ,Dai Jinhai ,Gui Xianzhou. QoS guarantee for stochastic real-time systems [A]. BICSC '99 [C] ,Beijing ,China ,1999 ,10 :494 - 499.

[7] 彭良智,戴金海,桂先洲,单懿. 强实时系统静态优先级调度的可调度性分析 [J]. 计算机工程与应用,1999 ,35(12) :13 - 16.

作者简介:

彭良智 1973 年生,1994 年和 1997 年分别获得国防科技大学控制理论与控制工程专业和模式识别与智能控制专业学士与硕士学位. 目前正攻读系统建模与仿真分析专业博士学位. 已在国内外发表学术论文 10 多篇. 主要研究方向为系统仿真、实时控制、实时系统及应用、图形图像处理.

戴金海 1943 年生,教授、博士生导师. 中国计算机用户协会仿真计算机分会副理事长. 1967 年毕业于哈尔滨工业大学工程力学系. 曾留学美国伊利诺伊大学计算机科学系. 主要从事系统建模与仿真、计算机实时控制与实时系统等研究工作. 曾两次获国家科技进步一等奖,三次获部级二等奖,在国内外学术刊物上发表论文多篇. 1992 年 10 月获国务院政府特殊津贴.