

# Heavy-Traffic Limits for Waiting Times in the $G_t / H_2^* / s_t + GI$ Queuing Model

Xin Niu, Jianmin Liu, Qingqing Wang

School of Science, Chang'an University, Xi'an Shaanxi  
Email: 1043614593@qq.com, jmliu@chd.ed.cn, 1027060533@qq.com

Received: Nov. 2<sup>nd</sup>, 2019; accepted: Nov. 20<sup>th</sup>, 2019; published: Nov. 27<sup>th</sup>, 2019

---

## Abstract

This article considered the queuing model, having time-varying arrival rate and staffing,  $H_2^*$  service times and customer abandonment according to a general probability distribution, analyzed a sequence of  $G_t / H_2^* / s_t + GI$  queue model, expressed the head of line (HOL) waiting time and obtained the Functional weak law of large numbers of HOL waiting time, and proved it.

## Keywords

Time-Varying, Waiting Time, Busy Interval, Heavy-Traffic Limit, Tightness

---

# $G_t / H_2^* / s_t + GI$ 队列模型等待时间的高负荷极限

牛 鑫, 刘建民, 王青青

长安大学理学院, 陕西 西安  
Email: 1043614593@qq.com, jmliu@chd.ed.cn, 1027060533@qq.com

收稿日期: 2019年11月2日; 录用日期: 2019年11月20日; 发布日期: 2019年11月27日

---

## 摘 要

考虑的队列模型为具有随时间变化的到达率和服务台个数, 服从  $H_2^*$  服务时间分布, 顾客放弃服从一般的概率分布。通过分析  $G_t / H_2^* / s_t + GI$  队列模型的一个序列, 对队首等待时间进行表示, 建立队首等待

时间的泛函弱大数定律, 并给出证明。

## 关键词

随时间变化, 等待时间, 忙期间隔, 高负荷极限, 胎紧性

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着排队系统理论体系的发展, 大量学者对不同的排队模型进行了研究。Liu, Y.N.和 Whitt, W. [1]建立了在服务时间为指数分布的情况下, 多服务台队列模型的高负荷极限。Whitt, W. [2]考虑  $H_2^*$  服务时间分布, 建立了  $G/GI/n/m$  队列模型中队长、等待时间及溢出随机过程的高负荷极限。Pang, G.D., Whitt, W. [3]通过建立双参数随机过程的高负荷极限, 得到  $G/GI/\infty$  队列的高负荷逼近。Liu, Y.N.和 Whitt, W. [4]用确定的流体模型来逼近多服务台  $G_t/GI/s_t+GI$  队列模型, 并确定与时间相关的性能函数。Liu, Y.N.和 Whitt, W. [5]建立了  $G_t/GI/s_t+GI$  在高负荷下的泛函弱大数定律。Talreja, R., Whitt, W. [6]通过建立二维的 Puhalskii 首达时间的不变性定理, 得到带有放弃的多服务台队列的等待时间的高负荷极限。刘建民 [7]通过引入奥伦斯坦-乌伦贝克过程和布朗运动, 在高负荷条件下, 得到  $GI/GI/m$  队列中负荷过程和队长过程的收敛极限。Whitt, W. [8]建立了  $G/GI/n/m$  队列模型中队长的扩散逼近, 通过用扩散过程的稳态分布得到队列模型测度指标的近似值, 重点研究了稳态延迟率。谢瑞金 [9]主要研究了  $G/G/n/m_n+G$  队列模型中等待时间在高负荷下的随机过程极限。Halfin, S., Whitt, W. [10]证明了关于  $s$  个服务台队列的两种不同的高负荷极限定理。Puhalskii, A.A. [11]研究了在高负荷条件下, 带有顾客放弃的具有周期性的多服务台队列中顾客数的极限。根据文献[1], 建立服务时间分布  $H_2^*$  服务时间分布的  $G_t/H_2^*/s_t+GI$  队列模型, 运用随机过程极限[12]及概率测度收敛[13]的相关知识, 对  $G_t/H_2^*/s_t+GI$  队列模型中队首等待时间在高负荷期间的高负荷极限进行证明。

## 2. $G_t/H_2^*/s_t+GI$ 队列模型描述

$G_t/H_2^*/s_t+GI$  队列模型, 具有随时间变化的到达率和服务台数, 一般到达过程且满足泛函中心极限定理(FCLT),  $H_2^*$  服务时间分布( $H_2^*$  服务时间分布是由以概率为  $p$  的指数分布和以概率为  $1-p$  的零点集混合而成的), 顾客放弃服从一般的概率分布, 有无穷等待空间且服从先到先服务的服务原则。

令  $S$  表示与  $H_2^*$  服务时间分布相关的量, 设  $H_2^*$  服务时间分布与  $n$  是独立的,  $H_2^*$  服务时间分布的均值为  $\frac{1}{\mu}$ ,  $0 < \mu < \infty$ , 则负荷强度是关于  $n$  的函数, 即  $\rho_n = \frac{\lambda_n}{\mu n}$ , 设  $H_2^*$  服务时间分布中指数部分的均值为  $\frac{1}{\nu}$ , 则有  $\frac{1}{\mu} = p \frac{1}{\nu}$ 。设  $\mathbb{D} \equiv D(I)$  表示在区间  $I \in \mathbb{R}$  上的左极限存在的右连续实值函数空间, 属于  $J_1$  拓扑, 即在区间  $I$  上的所有紧子区间中, 都可以由连续极限得到一致收敛,  $\Rightarrow$  表示依分布收敛。

考虑  $G_t/H_2^*/s_t+GI$  队列模型的一个序列, 指数为  $n$ 。在第  $n$  个模型中, 具有一般到达过程, 随时间变化的到达率为  $\lambda_n(t) \equiv n\lambda(t)$ , 独立同分布(i.i.d.)的  $H_2^*$  服务时间分布其累积分布函数(cdf)为  $G(t) \equiv 1 - e^{-\rho t}$ , 随时间变化的服务台数为  $s_n(t) \equiv \lceil ns(t) \rceil$ , 其中  $\lceil x \rceil$  表示大于等于  $x$  的最小整数, 顾客

在队列中等待进入服务的耐心时间是 i.i.d.且具有一般的 *cdf*  $F$ , 假设  $F$  是可微的, 其概率密度函数(pdf) 为  $f$ , 即  $\forall x$ , 有  $F^c(x) > 0$  和  $f(x) > 0$ 。

设  $N_n(t)$  表示在  $[0, t]$  时间间隔内到达的顾客数, 假设到达过程的序列  $\{N_n\}$  满足具有随时间变换布朗极限的 FCLT, 即当  $n \rightarrow \infty$  时, 有

$$\hat{N}_n(t) \equiv \frac{N_n(t) - n\Lambda(t)}{\sqrt{n}} \Rightarrow \hat{N}(t) \equiv c_\lambda B_\lambda(\Lambda(t)) \quad (1)$$

其中  $B_\lambda$  为标准的布朗运动(下标  $\lambda$  表示与到达过程相关),  $\Lambda(t)$  表示在  $[0, t]$  时间间隔内的总到达率, 即  $\Lambda(t) \equiv \int_0^t \lambda(s) ds$ ,  $c_\lambda^2$  是关于到达过程变量的参数。

设  $X_n(t)$  表示系统中的随时间变化的总顾客数(包括在服务中和队列中等待的顾客数), 有

$$X_n(t) \approx nX(t) + \sqrt{n}\hat{X}(t) = N(nX(t), n\sigma_X^2(t)) \quad (2)$$

其中,  $N(m, \sigma^2)$  表示均值为  $m$ , 方差为  $\sigma^2$  的高斯分布,  $X(t)$  为确定的流体逼近,  $\hat{X}(t)$  是均值为 0, 方差为  $\sigma_X^2(t) \equiv \text{Var}(\hat{X}(t))$  的高斯过程。

设  $B_n(t)$  表示  $t$  时刻在服务中的总顾客数;  $Q_n(t)$  表示  $t$  时刻在队列中的总顾客数;

$X_n(t) \equiv B_n(t) + Q_n(t)$  表示  $t$  时刻在系统中的总顾客数;

$W_n(t)$  表示  $t$  时刻的队首等待时间(HWT), 即  $t$  时刻在队列最前面顾客的等待时间;

$V_n(t)$  表示  $t$  时刻的虚等待时间(PWT), 即  $t$  时刻新到达顾客后不放弃的等待时间;

$A_n(t)$  表示在  $[0, t]$  时间间隔内放弃的顾客数;

$D_n(t)$  表示在  $[0, t]$  时间间隔内服务完成后离开的顾客数;

$F(x)$  表示进入队列但没有进入服务, 按时间  $x$  放弃的比例。

假设服务台变化函数  $s(t)$  是可行的, 即当  $s(t)$  减小时, 不会有顾客被强制退出服务系统。

由流量守恒可得, 当  $t \geq 0$  时, 有

$$A_n(t) = X_n(0) + N_n(t) - D_n(t) - X_n(t) \quad (3)$$

### 3. 主要结果及证明

假设在该队列模型中,  $H_2^*$  服务时间分布中指数服务分布的部分称为“忙期”, 零点集分布部分称为“闲期”。假设系统中“忙期”与“闲期”是交替进行的。

假设系统是从忙期开始, 为方便研究, 给定一个特殊的初始条件, 即所有服务台都是忙的且队列为空, 即系统中的初始人数为  $X(0) = B(0) = s(0)$ , 且有  $Q(0) = 0$ 。

假设在忙期间隔  $[\tau_1, \tau_2]$  内, 进入服务的速率大于 0, 顾客进入服务的速率是由服务台变空闲的量决定的, 则当  $\tau_1 \leq t \leq \tau_2$ , 有

$$b(t, 0) = s(t) p\mu + s'(t) \geq b_{\text{inf}} > 0 \quad (4)$$

在忙期间隔的服务速率为:  $b(t, x) = b(t-x, 0)G^c(x)1_{\{x \leq t\}} + b(0, x-t)\frac{G^c(x)}{G^c(x-t)}1_{\{x > t\}}$ 。

其中  $G^c(x) = 1 - G(x) = e^{-p\mu x}$ ,  $b(t, 0) = s(t) p\mu + s'(t)$  表示在  $t$  时刻进入服务的速率,  $b(0, x)$  表示初始进入服务的速率。

在忙期间隔内, 队长主要在假设没有人进入服务的情况下通过队列中人数密度  $\tilde{q}(t, x)$  来分析, 假设队列初始时空队列为空, 则  $\tilde{q}(t, x) = \lambda(t-x)F^c(x)$ ,  $x \leq t$ 。

由文献[4]推论 2 可知, 队长的密度为

$$q(t, x) = \tilde{q}(t, x) 1_{\{x \leq w(t)\}} = \lambda(t-x) F^c(x) 1_{\{x \leq w(t)\}} \quad (5)$$

**引理 1** (文献[4]) 考虑一个高负荷间隔  $[0, T]$ , 如果  $0 \leq w(0) < \infty$ ,  $\lambda_{\inf}(t) \equiv \inf_{0 \leq u \leq t} \{\lambda(u)\} > 0$ , 且当  $w(0) > 0$  时,  $q_{\inf}(0) \equiv \inf_{0 \leq u \leq w(0)} \{q(0, u)\} > 0$  成立, 则队首等待时间  $w$  是常微分程

$$w'(t+) = \Psi(t, w(t)) \equiv 1 - \frac{b(t+, 0)}{\tilde{q}(t, w(t)-)} \quad (6)$$

关于任何初值  $w(0)$  在  $[0, T]$  上的唯一解。

此外,  $w$  在  $[0, T]$  上是利普希茨连续的, 即  $\forall t \geq 0, u \geq 0, t+u \leq T$ , 有  $w(t+u) \leq w(t)+u$ 。

$w$  是右可微的, 右导数为  $w'(t+)$  由式(6)给出。

$w$  也是左可微的, 左导数为

$$w'(t-) = \tilde{\Psi}(t, w(t)) \equiv 1 - \frac{b(t-, 0)}{\tilde{q}(t, w(t)+)} \quad (7)$$

由上可知, 在除了有限多个点  $t$  外,  $w$  是处处连续可微的。

**引理 2** (文献[4]) 考虑一个高负荷间隔  $[0, T]$ , 且  $0 \leq w(0) < \infty$ ,  $\lambda_{\inf}(t) \equiv \inf_{0 \leq u \leq t} \{\lambda(u)\} > 0$ , 且当  $w(0) > 0$  时,  $q_{\inf}(0) \equiv \inf_{0 \leq u \leq w(0)} \{q(0, u)\} > 0$  成立,  $w(0) = 0$ 。若  $\forall t \geq 0, \nu(t) < \infty$ , 则  $\nu$  满足

$$\nu(t-w(t)) = w(t) \text{ 或 } \nu(t) = w(t+\nu(t)), t \geq 0 \quad (8)$$

此外,  $\nu$  在点  $t$  是连续的, 当且仅当  $\exists \varepsilon > 0$ , 使得  $w(t+\nu(t)+\varepsilon) = w(t+\nu(t))+\varepsilon$ 。反之也成立, 当且仅当  $t+\nu(t) \leq u \leq t+\nu(t)+\varepsilon$ , 有  $b(u, 0) = 0$ , 若  $b(\cdot, 0) > 0$ , 则  $\nu$  是连续的。

由引理 1, 引理 2 可知, 在  $G_t / H_2^* / s_t + GI$  队列模型中 HWT  $w$  是下面常微分方程(ODE)的唯一解

$$w'(t) \equiv \frac{dw}{dt}(t) = 1 - \frac{b(t, 0)}{\tilde{q}(t, w(t))} = 1 - \frac{s(t) p \mu + s'(t)}{\lambda(t-w(t)) F^c(w(t))} \quad (9)$$

由上述假设, 可知上式中的分子与分母都是大于 0 的, 因此有

在忙期间,  $\forall t$ , 有  $-\infty < w'(t) < 1$

由式(9)的 ODE 可得

$$w(t) = \int_0^t \left( 1 - \frac{b(u, 0)}{\tilde{q}(u, w(u))} \right) du, t \geq 0 \quad (10)$$

由引理 2 可知, PWT  $\nu(t)$  是式(8)的唯一解, 由式(8)也可以解得 HWT  $w$ 。

由式(4)可知,  $\nu$  是连续函数, 事实上,  $w$  和  $\nu$  除在有限多个点外都是可微的, 由上式可得其导数为:

$$\nu'(t-w(t)) = \frac{w'(t)}{1-w'(t)} \text{ 或 } \nu'(t) = \frac{w'(t+\nu(t))}{1-w'(t+\nu(t))} \quad (11)$$

由式(4)可知, 上式是有界的。

在忙期间隔内, 服务台都是满的, 则在  $[0, t]$  时间间隔内完成服务离开的顾客数  $D(t) = S(t) p \mu$ , 其中  $S(t) \equiv \int_0^t s(u) du$ 。

在  $[0, t]$  时间间隔内放弃的顾客数为  $A(t) = \int_0^t \alpha(s) ds$ , 其中

$$\alpha(s) = \int_0^\infty Q(s, x) h_F(x) dx \quad (12)$$

$h_F(x) \equiv \frac{f(x)}{F^c(x)}$  是关于累积分布函数  $F$  的风险率函数, 因为  $f$  是  $\mathbb{D}$  中的元素, 对所有  $x$ , 有  $F^c(x) > 0$ ,

则  $h_F(x)$  是有界的。

因为假设系统开始时为忙期, 则忙期间隔为  $[\tau_{2i}, \tau_{2i+1}]$ ,  $i \geq 0$ , 闲期间隔为  $[\tau_{2i+1}, \tau_{2i+2}]$ ,  $i \geq 0$ , 有限多个这样的间隔可以覆盖整个区间  $[0, T]$ 。现在用递归的方法考虑这些间隔, 为方便表示, 下文中记作  $[0, \tau]$ 。

先考虑一个忙期间隔  $[0, \tau]$ , 假设当  $n \rightarrow \infty$  时,  $\bar{X}_n(0) \Rightarrow X(0)$ , 有  $X(0) = s(0)$ 。

通过分析进入服务的顾客数来研究队首等待时间。

设  $E_n(t)$  表示在  $[0, t]$  时间间隔内进入服务的顾客数,  $\forall t, 0 < t < \tau$ , 有  $\Lambda(t) > D(t)$ 。则进入服务的人数与因完成服务而离开造成服务台空闲和服务台的变化量之和是渐进相等的。

即, 当  $n \rightarrow \infty$  时,

$$\sup_{0 < t < \tau} \left\{ \left| E_n(t) - (D_n(t) + \lceil ns(t) \rceil - \lceil ns(0) \rceil) \right| \right\} = o(\sqrt{n}) \quad (13)$$

对  $E_n(t)$  进行刻画, 有

$$\bar{E}_n(t) \equiv \frac{E_n(t)}{n}, \quad \hat{E}_n(t) \equiv \sqrt{n}(\bar{E}_n(t) - E(t)) \quad (14)$$

在  $\mathbb{D}$  中, 当  $n \rightarrow \infty$  时, 有

$$\bar{E}_n(t) \Rightarrow E(t), \quad \hat{E}_n(t) \Rightarrow \hat{E}(t) \quad (15)$$

其中  $E(t) \equiv D(t) + s(t) - s(0)$ ,  $\hat{E}(t) = \hat{D}(t) = B_s(D(t))$ 。

关于  $E_n(t)$  还有另一种表示方式: 因为进入服务的顾客是等待时间最长的顾客,

$$Q_n(t) = \tilde{Q}_n(t, W_n(t)) = \sum_{i=N_n(t-W_n(t)-)+1}^{N_n(t)} 1(\tau_i^n + \eta_i > t) \text{ 表示在 } [0, t] \text{ 时间间隔内进入服务的顾客数} \quad (16)$$

其中  $\tau_i^n$  表示在模型  $n$  中第  $i$  个顾客到达的时间;  $\eta_i$  表示第  $i$  个顾客的等待时间。

设  $E_n(t, \varepsilon) \equiv E_n(t + \varepsilon) - E_n(t)$ , 若时间间隔足够小的时候, 放弃的人数可以忽略。

设  $I_n(t, \varepsilon)$  表示在  $[t, t + \varepsilon]$  内从队列前面离开的顾客数,  $A_n^l(t, \varepsilon)$  表示  $I_n(t, \varepsilon)$  中放弃的顾客数

$$E_n(t, \varepsilon) = I_n(t, \varepsilon) - A_n^l(t, \varepsilon) \quad (17)$$

其中

$$I_n(t, \varepsilon) = \sum_{i=N_n(t-W_n(t)-)+1}^{N_n(t+\varepsilon-W_n(t+\varepsilon))} 1(\tau_i^n + \eta_i > t) \quad (18)$$

因为队列中顾客的等待时间是以速率 1 增加的, 则有  $W_n(t + \varepsilon) \leq W_n(t) + \varepsilon$ 。

下证  $A_n^l(t, \varepsilon)$  是可以忽略的。

令  $J_n(t, \varepsilon)$  表示在  $t$  时刻在系统中的顾客, 在  $[t, t + \varepsilon]$  时间间隔内没有进入服务, 在  $t + \varepsilon$  时刻之前会

放弃的顾客数。

可知,  $I_n(t, \varepsilon) - J_n(t, \varepsilon)$  的顾客数一定会进入服务, 因为在  $t + \varepsilon$  时刻之前没有放弃。

首先考虑到达时间  $\tau_i^n$ ,  $0 < \tau_i^n < \tau$ , 则

$$\begin{aligned} J_n(t, \varepsilon) &\leq I_n(t, \varepsilon) \sup \left\{ P(t + \tau_i^n + \eta_i \leq t + \varepsilon \mid \tau < \tau_i^n + \eta_i) \right\} \\ &\leq I_n(t, \varepsilon) \sup_{0 \leq t \leq \tau} \left\{ F^c(t) - F^c(t + \varepsilon) \right\} \\ &\leq I_n(t, \varepsilon) \|f\|_r \varepsilon + o(\varepsilon) \end{aligned}$$

则

$$A_n^I(t, \varepsilon) \leq J_n(t, \varepsilon) \leq K \varepsilon I^n(t, \varepsilon), \quad K = \|f\|_r \quad (19)$$

由式(13)和(16)等价, 可以对队首等待时间  $W_n(t)$  进行描述。

如果系统一直处于忙期, 可以使用无穷服务台模型, 如式(13)中, 其误差可以忽略。

我们可以得到式(18)的另一种表达式

$$I_n(t, \varepsilon) = n \int_{t-W_n(t)}^{t+\varepsilon-W_n(t+\varepsilon)} \int_0^\infty 1(s+x > t) d\bar{K}_n(\bar{N}_n(s), x) \quad (20)$$

其中  $\bar{K}_n(t, y) \equiv \frac{1}{n} \sum_{i=1}^{\lfloor mt \rfloor} 1(\eta_i \leq y)$ ,  $t \geq 0$ ,  $y \geq 0$ 。

根据文献[3]可知

$$I_n(t, \varepsilon) = I_{n,1}(t, \varepsilon) + I_{n,2}(t, \varepsilon) + I_{n,3}(t, \varepsilon) \quad (21)$$

其中  $I_{n,1}(t, \varepsilon) = \sqrt{n} \int_{t-W_n(t)}^{t+\varepsilon-W_n(t+\varepsilon)} F^c(t-s) d\hat{N}_n(s)$ 。

$$I_{n,2}(t, \varepsilon) = \sqrt{n} \int_{t-W_n(t)}^{t+\varepsilon-W_n(t+\varepsilon)} \int_0^\infty 1(s+x > t) d\hat{R}_n(s, x) \quad (22)$$

$$I_{n,3}(t, \varepsilon) = n \int_{t-W_n(t)}^{t+\varepsilon-W_n(t+\varepsilon)} F^c(t-s) \lambda(s) ds$$

$$\hat{R}_n(s, x) \equiv \sqrt{n} \bar{K}_n(\bar{N}_n(s), x) - \hat{N}_n(s) F(x) - \sqrt{n} \Lambda(ts) F(x)$$

**定理(等待时间的泛函弱大数定律)**

考虑一个忙期间隔  $[0, \tau]$ 。若  $\bar{X}_n(0) \Rightarrow X(0)$  成立, 且  $X(0) = s(0)$ , 则在  $\mathbb{D}([0, \tau])$  中, 当  $n \rightarrow \infty$  时, 有

$$W_n \Rightarrow w \quad (23)$$

其中  $w$  为式(9)中连续确定性流体模型函数的极限。

**证明:** 为方便证明, 我们考虑特殊的初始条件, 即所有服务台都是忙的且队列为空。可知在忙期间隔离去过程与速率为  $s(t) p\mu$  的非齐次泊松过程是渐进相等的。要证  $W_n \Rightarrow w$ 。通过利用判别紧性的方法(参考文献[12], 第 11.6 节)来证  $W$  的泛函弱大数定律, 我们需要先证明序列  $\{W_n\}$  在  $\mathbb{D}([0, \tau])$  上是胎紧的, 然后证明每个收敛子序列的极限相等。

先证  $\{W_n\}$  的胎紧性

首先序列  $\{W_n\}$  是有界的, 因为  $W_n(t) \geq 0$  且  $W_n(t)$  最大是以速率 1 增加的。所考虑的忙期间隔是在一个较大的有限间隔  $[0, \tau^*]$  中, 因为系统假设是从忙期间隔开始且队列为空, 则  $W_n(t) \leq \tau^*$ 。

在闲期间隔内,  $W_n(0) \Rightarrow 0$  仍然成立, 所以  $\limsup_{n \rightarrow \infty} W_n(t) \leq \tau$ 。

因为对所有的  $t$ ,  $W_n(t+\delta) - W_n(t) \leq \delta$ , 则  $W_n(t)$  的连续模有界。

设  $\bar{I}_{n,j} \equiv \frac{I_{n,j}}{n}$ , 在  $\mathbb{D}([0, \tau])$  中,  $\bar{I}_{n,1} \Rightarrow 0e$ ,  $\bar{I}_{n,2} \Rightarrow 0e$  则

$$\|\bar{I}_n - \bar{I}_{n,3}\|_r \Rightarrow 0. \quad (24)$$

由式(11)可知,  $\bar{I}_{n,3}(t, \delta) - \bar{A}_n^I(t, \delta) \Rightarrow E(t, \delta)$ , 则  $\bar{A}_n^I(t, \delta)$  是可以忽略的。

由对  $\lambda$  和  $F$  的假设, 可知当  $c \equiv F^c(\tau) \lambda_{\text{mf}}(\tau) > 0$  时,  $I_{n,3}$  的被积函数是有界的。

有

$$\frac{\bar{I}_{n,3}(t, \delta)}{c} \geq W_n(t) - W_n(t+\delta) + \delta \quad (25)$$

所以  $W_n(t) - W_n(t+\delta) \leq \frac{\bar{I}_{n,3}(t, \delta)}{c} - \delta \leq \frac{\bar{E}_n(t, \delta)}{c}$ 。

则有

$$\limsup_{n \rightarrow \infty} \{W_n(t) - W_n(t+\delta)\} \leq \frac{D(t, \delta) + s(t, \delta)}{c} \equiv K\delta, \quad K \text{ 为常数。} \quad (26)$$

即序列  $\{W_n(t): 0 \leq t \leq \tau\}$  是胎紧的。

再说明  $\{W_n\}$  收敛子序列的极限, 由胎紧性可知每个序列都有一个收敛的子序列, 则要证  $W_n \Rightarrow w$ , 现在只需证  $\{W_n\}$  的每个收敛子序列的极限为  $w$ 。即证  $w$  满足下式

$$w'(t) \equiv \frac{dw}{dt}(t) = 1 - \frac{b(t, 0)}{\bar{q}(t, w(t))} = 1 - \frac{s(t) p\mu + s'(t)}{\lambda(t - w(t)) F^c(w(t))} \quad \text{或} \quad w(t) = \int_0^t \left( 1 - \frac{b(u, 0)}{\bar{q}(u, w(u))} \right) du$$

因为  $w$  是上面方程的唯一解。

由式(13)和(16)可得

$$\bar{E}_n(t, \varepsilon) \Rightarrow E(t, \varepsilon) = \int_t^{t+\varepsilon} b(s, 0) ds, \quad n \rightarrow \infty \quad (27)$$

在上式中, 当  $\varepsilon \rightarrow \infty$  时, 其极限接近于  $b(t, 0) = s(t) p\mu + s'(t)$ 。

我们还需要考虑从队列中离开的顾客数。  $\bar{I}_n(t, \varepsilon)$  与  $\bar{I}_{n,3}(t, \varepsilon)$  是渐进相等的。

对  $\bar{I}_{n,3}(t, \varepsilon)$  应用  $W_n \Rightarrow W$  和连续映射定理可知, 在  $\mathbb{D}([0, \tau])$  中,

$$\bar{I}_n(t, \varepsilon) \Rightarrow I(t, \varepsilon) \equiv \int_{t-W(t)}^{t+\varepsilon-W(t+\varepsilon)} F^c(t-s) \lambda(s) ds \quad (28)$$

也就是说, 一旦我们知道极限  $W$  就可以确定  $I(t, \varepsilon)$  的极限。

由上可得,

$$\bar{A}_n^I(t, \varepsilon) = \bar{E}_n(t, \varepsilon) - \bar{I}_n(t, \varepsilon) \Rightarrow A^I(t, \varepsilon) \equiv E(t, \varepsilon) - I(t, \varepsilon) \quad (29)$$

在给定极限  $W$  的条件下,  $I(t, \varepsilon)$  和  $A^I(t, \varepsilon)$  极限可以确定。

则

$$E(t, \varepsilon) = I(t, \varepsilon) + A^I(t, \varepsilon) = \int_{t-W(t)}^{t+\varepsilon-W(t+\varepsilon)} F^c(t-s) \lambda(s) ds + A^I(t, \varepsilon) \quad (30)$$



由式(19)可知, 当  $\varepsilon$  任意小时,  $A^l(t, \varepsilon)$  可以忽略。

则

$$E(t, \varepsilon) = b(t, 0)\varepsilon + o(\varepsilon) = F^c(t - W(t))\lambda(W(t))(1 - W'(t)) + o(\varepsilon) \quad (31)$$

由引理 1 可知  $W$  满足式(9), 因为常微分方程有唯一解, 则  $P(W = w) = 1$ 。

因为以上结论对每个收敛子序列的极限都成立, 则  $W_n \Rightarrow w$ 。

因为  $W \Rightarrow w$ , 可以确定  $I(t, \varepsilon)$  和  $A^l(t, \varepsilon)$  的极限, 即用  $w(t)$  代替  $W(t)$ 。

即  $E(t)$  可以表示为

$$E(t) \equiv E(0, t) \text{ 或 } E(t) = D(t) + s(t) - s(0) = \int_0^{t-w(t)} F^c(t-s)\lambda(s)(1-w'(s))ds. \quad \square$$

#### 4. 结束语

本文是在特殊初始条件下对  $G_i / H_2^* / s_i + GI$  队列模型进行研究。通过将两种不同表达式等价, 对队首等待时间进行表示。先证明队首等待时间序列的胎紧性, 再说明每个收敛子序列收敛到同一个极限来证明队首等待时间的泛函弱大数定律。

#### 参考文献

- [1] Liu, Y. and Whitt, W. (2014) Many-Server Heavy-Traffic Limit for Queues with Time-Varying Parameters. *Annals of Applied Probability*, **24**, 378-421. <https://doi.org/10.1214/13-AAP927>
- [2] Whitt, W. (2005) Heavy-Traffic Limits for the  $G / H_2^* / n / m$  Queue. *Mathematics of Operations Research*, **30**, 1-27. <https://doi.org/10.1287/moor.1040.0119>
- [3] Pang, G. and Whitt, W. (2010) Two-Parameter Heavy-Traffic Limits for Infinite-Server Queues. *Queueing Systems*, **65**, 325-364. <https://doi.org/10.1007/s11134-010-9184-z>
- [4] Liu, Y. and Whitt, W. (2012) The  $G_i / GI / s_i + GI$  Many-Server Fluid Queue. *Queueing Systems*, **71**, 405-444. <https://doi.org/10.1007/s11134-012-9291-0>
- [5] Liu, Y. and Whitt, W. (2012) A Many-Server Fluid Limit for the  $G_i / GI / s_i + GI$  Queueing Model Experiencing Periods of Overloading. *Operations Research Letters*, **40**, 307-312. <https://doi.org/10.1016/j.orl.2012.05.010>
- [6] Talreja, R. and Whitt, W. (2009) Heavy-Traffic Limits for Waiting Times in Many-Server Queues with Abandonment. *Annals of Applied Probability*, **19**, 2137-2175. <https://doi.org/10.1214/09-AAP606>
- [7] 刘建民. 高负荷下带有放弃的  $GI / GI / m$  队列[J]. 工程数学学报, 2008, 25(2): 245-252.
- [8] Whitt, W. (2004) A Diffusion Approximation for the  $G / GI / n / m$  Queue. *Operations Research*, **52**, 922-941. <https://doi.org/10.1287/opre.1040.0136>
- [9] 解瑞金. 高负荷下带有放弃的排队系统的等待时间[D]: [硕士学位论文]. 西安: 长安大学, 2010.
- [10] Halfin, S. and Whitt, W. (1981) Heavy-Traffic Limits for Queues with Many Exponential Servers. <https://doi.org/10.1287/opre.29.3.567>  
<http://www.columbia.edu/~ww2040/allpapers.html>
- [11] Puhalskii, A.A. (2013) On the  $M_i / M_i / k_i + M_i$  Queue in Heavy Traffic. *Mathematical Methods of Operations Research*, **78**, 119-148. <https://doi.org/10.1007/s00186-013-0435-8>
- [12] Whitt, W. (2002) *Stochastic-Process Limits*. Springer, New York. <https://doi.org/10.1007/b97479>
- [13] Billingsley, P. (1999) *Convergence of Probability Measures*. 2nd Edition, Wiley, New York. <https://doi.org/10.1002/9780470316962>