Algorithms for Extremal Queues in GI/GI/1

Yan Chen Columbia University, IEOR Department

Joint work with Ward Whitt Columbia University, IEOR Department • Extremal queues in *GI/GI/1*: *F*₀/*G_u/1* (Chen and Whitt (2018)) (a special two-point inter-arrival and service times queue)

- Extremal queues in *GI/GI/1*: *F*₀/*G_u/1* (Chen and Whitt (2018)) (a special two-point inter-arrival and service times queue)
- Research Question: How to obtain the steady-state mean waiting time E[W(F₀/G_u/1)] (tight upper bound) ?

- Extremal queues in *GI/GI/1*: *F*₀/*G_u/1* (Chen and Whitt (2018)) (a special two-point inter-arrival and service times queue)
- Research Question: How to obtain the steady-state mean waiting time E[W(F₀/G_u/1)] (tight upper bound) ?
- GOAL: closed form solution of $E[W(F_0/G_u/1)]$?

- Extremal queues in *GI/GI/1*: *F*₀/*G_u/1* (Chen and Whitt (2018)) (a special two-point inter-arrival and service times queue)
- Research Question: How to obtain the steady-state mean waiting time E[W(F₀/G_u/1)] (tight upper bound) ?
- GOAL: closed form solution of $E[W(F_0/G_u/1)]$?
- GOAL: effective algoritms to estimate $E[W(F_0/G_u/1)]$?

• Steady-state Waiting Time ($W_0 = 0$):

$$W \stackrel{\mathrm{d}}{=} [W + V - U]^+.$$

$$w: \mathcal{P}_{a,2}(M_a) \times \mathcal{P}_{s,2}(M_s) \to \mathbb{R}, w(F,G) \equiv E[W(F,G)]$$

where $0 < \rho < 1$ and

• Mean Waiting Time:

$$E[W] = \sum_{k=1}^{\infty} \frac{E[S_k^+]}{k} < \infty.$$

 $S_k \equiv X_1 + \cdots + X_k$ and $X_k \equiv V_k - U_k$, $k \ge 1$.

 $(1, c_a^2 + 1)$ for interarrival F and $(\rho, \rho^2(c_s^2 + 1))$ for service G.

• Kingman (1962) bound:

$$E[W(F,G)] \leq \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)}.$$

 $(1, c_s^2 + 1)$ for interarrival F and $(\rho, \rho^2(c_s^2 + 1))$ for service G.

• Kingman (1962) bound:

$$E[W(F,G)] \leq \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)}.$$

• Daley (1977) bound:

$$E[W(F,G)] \leq \frac{\rho^2([(2-\rho)c_a^2/\rho] + c_s^2)}{2(1-\rho)}.$$

 $(1, c_s^2 + 1)$ for interarrival F and $(\rho, \rho^2(c_s^2 + 1))$ for service G.

• Kingman (1962) bound:

$$E[W(F,G)] \leq \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)}.$$

• Daley (1977) bound:

$$E[W(F,G)] \leq \frac{\rho^2([(2-\rho)c_a^2/\rho] + c_s^2)}{2(1-\rho)}.$$

• Heavy Traffic Approximation:

$$E[W(M,G)] = \frac{\rho^2(1+c_s^2)}{2(1-\rho)}.$$

 $(1, c_s^2 + 1)$ for interarrival F and $(\rho, \rho^2(c_s^2 + 1))$ for service G.

• Kingman (1962) bound:

$$E[W(F,G)] \leq \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)}.$$

• Daley (1977) bound:

$$E[W(F,G)] \leq rac{
ho^2([(2-
ho)c_a^2/
ho]+c_s^2)}{2(1-
ho)}.$$

• Heavy Traffic Approximation:

$$E[W(M,G)] = \frac{\rho^2(1+c_s^2)}{2(1-\rho)}.$$

• Ott (1987) bound:

$$E[W(D,A_3)] = \frac{\rho((1+c_s^2)\rho - 1)^+}{2(1-\rho)}$$

• $F: c_a^2/(c_a^2 + (b_a - 1)^2)$ at b_a , $(b_a - 1)^2/(c_a^2 + (b_a - 1)^2)$ at $1 - c_a^2/(b_a - 1)$ where $1 + c_a^2 \le b_a \le M_a$.

- $F: c_a^2/(c_a^2 + (b_a 1)^2)$ at b_a , $(b_a 1)^2/(c_a^2 + (b_a 1)^2)$ at $1 c_a^2/(b_a 1)$ where $1 + c_a^2 \le b_a \le M_a$.
- G: $c_s^2/(c_s^2 + (b_s 1)^2)$ at ρb_s , $(b_s 1)^2/(c_s^2 + (b_s 1)^2)$ at $\rho(1 c_s^2/(b_s 1))$ where $1 + c_s^2 \le b_s \le M_s$.

•
$$F: c_a^2/(c_a^2 + (b_a - 1)^2)$$
 at $b_a, (b_a - 1)^2/(c_a^2 + (b_a - 1)^2)$ at $1 - c_a^2/(b_a - 1)$ where $1 + c_a^2 \le b_a \le M_a$.

- G: $c_s^2/(c_s^2 + (b_s 1)^2)$ at ρb_s , $(b_s 1)^2/(c_s^2 + (b_s 1)^2)$ at $\rho(1 c_s^2/(b_s 1))$ where $1 + c_s^2 \le b_s \le M_s$.
- $F = F_0$ for $b_a = 1 + c_a^2$ and $F = F_u$ when $b_a = M_a$.

•
$$F: c_a^2/(c_a^2 + (b_a - 1)^2)$$
 at $b_a, (b_a - 1)^2/(c_a^2 + (b_a - 1)^2)$ at $1 - c_a^2/(b_a - 1)$ where $1 + c_a^2 \le b_a \le M_a$.

- G: $c_s^2/(c_s^2 + (b_s 1)^2)$ at ρb_s , $(b_s 1)^2/(c_s^2 + (b_s 1)^2)$ at $\rho(1 c_s^2/(b_s 1))$ where $1 + c_s^2 \le b_s \le M_s$.
- $F = F_0$ for $b_a = 1 + c_a^2$ and $F = F_u$ when $b_a = M_a$.
- $G = G_0$ for $b_s = 1 + c_s^2$ and $G = G_u$ when $b_s = M_s$.

Reduction

• Three-point reduction:

 $\sup_{F\in\mathcal{P}_{a,2}(M_a),G\in\mathcal{P}_{s,2}(M_s)} E[W(F,G)] = E[W(F^*,G^*)]$

where $F^* \in \mathcal{P}_{a,2,3}(M_a), G^* \in \mathcal{P}_{s,2,3}(M_s).$

Reduction

• Three-point reduction:

$$\sup_{F \in \mathcal{P}_{a,2}(M_a), G \in \mathcal{P}_{s,2}(M_s)} E[W(F,G)] = E[W(F^*,G^*)]$$

where $F^* \in \mathcal{P}_{a,2,3}(M_a), G^* \in \mathcal{P}_{s,2,3}(M_s).$

• Two-point reduction:

 $\sup_{F\in\mathcal{P}_{a,2,3}(M_a),G\in\mathcal{P}_{s,2,3}(M_s)}E[W(F,G)]=E[W(F_0,G_u)].$

Reduction

• Three-point reduction:

$$\sup_{F \in \mathcal{P}_{a,2}(M_a), G \in \mathcal{P}_{s,2}(M_s)} E[W(F,G)] = E[W(F^*,G^*)]$$

where $F^* \in \mathcal{P}_{a,2,3}(M_a), G^* \in \mathcal{P}_{s,2,3}(M_s)$.

• Two-point reduction:

$$\sup_{F\in\mathcal{P}_{\mathfrak{s},2,3}(M_{\mathfrak{s}}),G\in\mathcal{P}_{\mathfrak{s},2,3}(M_{\mathfrak{s}})} E[W(F,G)] = E[W(F_0,G_u)].$$

Research Question: define $E[W(F_0, G_{u^*})] \equiv \lim_{M_s \to \infty} E[W(F_0, G_u)]$, how to evaluate $E[W(F_0, G_{u^*})]$?

For any service-time V with cdf G having mean ρ and scv c_s^2 , the steady-state waiting time is distributed as (denoted by $\stackrel{d}{=}$)

$$W(F_0/G/1) \stackrel{d}{=} W(D(1/p)/RS(V,p)/1) + \sum_{k=1}^{N(p)-1} V_k$$

for N(p) and RS(V, p). Hence, the mean is

 $E[W(F_0/G/1)] = E[W(D(1/p)/RS(V,p)/1)] + (E[N(p)] - 1)E[V]$

- $= E[W(D(1/p)/RS(V,p)/1)] + \rho(1-p)/p$
- $= E[W(D(1/p)/RS(V,p)/1)] + \rho c_a^2.$

(the Daley decomposition) Consider the $F/G_u/1$ model with arbitrary interarrival-time cdf F and two-point service-time cdf G_u . Then

 $\lim_{M_s \to \infty} E[W(F/G_u/1)] = E[W(F/D(\rho)/1)] + \lim_{M_s \to \infty} E[W(D(1)/G_u/1)].$ = $E[W(F/D(\rho)/1)] + \frac{\rho^2 c_s^2}{2(1-\rho)}.$

(overall decomposition of the upper bound) For the GI/GI/1 model with extremal interarrival-time cdf F_0 and extremal service-time cdf G_{u^*} ,

$$E[W(F_0, G_{u^*})] \equiv \lim_{M_s \to \infty} E[W(F_0/G_u/1)]$$

= $E[W(D(1/p)/RS(D(\rho), p)/1)] + \rho c_a^2 + \frac{\rho^2 c_s^2}{2(1-\rho)}$

for $p \equiv 1/(c_a^2+1)$.

Numerical Algorithm

A new service time:

$$RS(V, p) \stackrel{\mathrm{d}}{=} \sum_{k=1}^{N(p)} V_k, \qquad (1)$$

where N(p) is a geometric random variable on the positive integers, having mean E[N(p)] = 1/p with $p = 1/(1 + c_a^2)$ and $\{V_k : k \ge 1\}$ is i.i.d. random variables distributed as a service time V.

Numerical Algorithm

A new service time:

$$RS(V,p) \stackrel{\mathrm{d}}{=} \sum_{k=1}^{N(p)} V_k, \qquad (1)$$

where N(p) is a geometric random variable on the positive integers, having mean E[N(p)] = 1/p with $p = 1/(1 + c_a^2)$ and $\{V_k : k \ge 1\}$ is i.i.d. random variables distributed as a service time V.

• Apply Inter-arrival Time Reduction:

$$E[W(F_0, G)] = E[W(D(1/p), RS(V, p))] + \rho c_a^2.$$

Numerical Algorithm

A new service time:

$$RS(V,p) \stackrel{\mathrm{d}}{=} \sum_{k=1}^{N(p)} V_k, \qquad (1)$$

where N(p) is a geometric random variable on the positive integers, having mean E[N(p)] = 1/p with $p = 1/(1 + c_a^2)$ and $\{V_k : k \ge 1\}$ is i.i.d. random variables distributed as a service time V.

• Apply Inter-arrival Time Reduction:

$$E[W(F_0, G)] = E[W(D(1/p), RS(V, p))] + \rho c_a^2.$$

• Apply Daley Decomposition:

 $E[W(F_0, G_{u^*})] = E[W(D(1/p), RS(D(\rho), p))] + const.$

Lemma (*NB* representation) For the $D(1/p)/RS(D(\rho), p)/1$ model,

$$S_k \stackrel{\mathrm{d}}{=}
ho(\mathsf{NB}(n,1-p)+n) - (n/p),$$

so that

$$E[W] = \sum_{n=1}^{\infty} \frac{E[(\rho(NB(n, 1-p) + n) - (n/p))^+]}{n}.$$

Lemma (*NB* representation) For the $D(1/p)/RS(D(\rho), p)/1$ model,

$$S_k \stackrel{\mathrm{d}}{=} \rho(\mathsf{NB}(n, 1-p) + n) - (n/p),$$

so that

$$E[W] = \sum_{n=1}^{\infty} \frac{E[(\rho(NB(n, 1-p) + n) - (n/p))^+]}{n}.$$

pmf: $P(NB(n, 1-p) = k) \equiv \left(\frac{(n+k-1)!}{k!(n-1)!}\right) p^n (1-p)^k$.

Lemma (*NB* representation) For the $D(1/p)/RS(D(\rho), p)/1$ model,

$$S_k \stackrel{\mathrm{d}}{=} \rho(\mathsf{NB}(n, 1-p) + n) - (n/p),$$

so that

$$E[W] = \sum_{n=1}^{\infty} \frac{E[(\rho(NB(n, 1-p) + n) - (n/p))^+]}{n}.$$

pmf: $P(NB(n, 1-p) = k) \equiv \left(\frac{(n+k-1)!}{k!(n-1)!}\right) p^n (1-p)^k$. recursion: $\mathbb{P}(NB = k) = \mathbb{P}(NB = k-1)(n+k-1)/k)(1-p)$. Algorithm 1 Basic Negative Binomial Recursion (k in outer loop)

1: for $k \in [K]$ do 2: $S(k) \leftarrow 0$, $nbpdf \leftarrow p(1-p)^k$ 3: for $n \in [n]$ do 4: $S(k) \leftarrow S(k) + nbpdf \max((n+k)\rho - n/p, 0)/n$ 5: $nbpdf \leftarrow nbpdf(\frac{n+k}{n})p$ 6: $E[W] \leftarrow E[W] + S(k)$

7: Output E[W]

Numerical Algorithm (Negative Binomial)

Refinement: Staring at mean and going up and down.

Algorithm 2 NB Recursion (Up and Down from the Mean)

1: for
$$n \in [1, N]$$
 do
2: $nbpdf(1, n(1-p)/p) \leftarrow 1$
3: for $k \in [m(n) - \alpha \sqrt{N}, m(n)]$ do
4: $nbpdf(1, k - 1) \leftarrow nbpdf(1, k)/(n + k - 1)(k)/(1 - p)$
5: for $k \in [m(n), m(n) + \alpha \sqrt{N} - 1]$ do
6: $nbpdf(1, k + 1) \leftarrow nbpdf(1, k)(n + k)/(k + 1)(1 - p)$
7: Normalize $nbpdf$ to obtain $\mathbb{P}(NB(n, 1 - p) = k)$
8: $S(n) \leftarrow \sum_{k} \mathbb{P}(NB(n, 1 - p) = k) \max((n + k)\rho - n/p, 0)$
9: $E[W] \leftarrow E[W] + S(n)/n$
10: Output $E[W]$

Table 1: Comparison of Two Approaches Generating Negative Binomial Probabilities

k	$n_1 = 10$	$n_2 = 10$	k	$n_1 = 100$	$n_2 = 100$	k	$n_1 = 1000$	$n_2 = 1000$
40	0.0279638	0.0279638	400	0.0089128	0.0089128	4000	0	0.0028207
41	0.0272818	0.0272818	401	0.0088906	0.0088906	4001	0	0.0028200
42	0.0265023	0.0265023	402	0.0088641	0.0088641	4002	0	0.0028192
43	0.0256394	0.0256394	403	0.0088333	0.0088333	4003	0	0.0028182
44	0.0247071	0.0247071	404	0.0087983	0.0087983	4004	0	0.0028170
45	0.0237188	0.0237188	405	0.0087592	0.0087592	4005	0	0.0028158
46	0.0226875	0.0226875	406	0.0087160	0.0087160	4006	0	0.0028144
47	0.0216256	0.0216256	407	0.0086689	0.0086689	4007	0	0.0028128
48	0.0205443	0.0205443	408	0.0086179	0.0086179	4008	0	0.0028111
49	0.0194542	0.0194542	409	0.0085631	0.0085631	4009	0	0.0028093
50	0.0183647	0.0183647	410	0.0085047	0.0085047	4010	0	0.0028074

(the idle-time representation) In the GI/GI/1 queue with cdf's F and G having parameter 4-tuple $(1, c_a^2, \rho, c_s^2)$,

$$E[W] \equiv E[W(F,G)] = \psi(1,c_a^2,\rho,c_s^2) - \phi(I),$$

where

$$\psi(1, c_a^2, \rho, c_s^2) \equiv \frac{E[(U-V)^2]}{2E[U-V]} = \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)} + \frac{1-\rho}{2}$$

and

$$\phi(I) \equiv \phi(F, G) = \frac{E[I^2]}{2E[I]} = E[I_e].$$

Corollary

(reduction to idle time) For the GI/GI/1 model with extremal interarrival-time cdf F_0 and extremal service-time cdf G_u ,

$$\begin{split} E[W(F_0, G_{u^*})] &\equiv \lim_{M_s \to \infty} E[W(F_0/G_u/1)] \\ &= \frac{c_a^2 + \rho^2 c_s^2}{2(1-\rho)} + \frac{1-\rho}{2} - \phi(I; 1, c_a^2, \rho, c_s^2), \end{split}$$

where I is the idle time in an $F_0/G_{u^*}/1$ queue or, equivalently, in a $F_0/D/1$ queue for an appropriate D.

Table 2: Performance of Algorithm with Different Truncation Levels

			Minh and Sorli Algorithm				
$\rho \setminus N$	2E+03	4E+03	8E+03	1.6E+04	2E+04	T = 1E + 07	95%CI
0.1	0.422229	0.422229	0.422229	0.422229	0.422229	0.422	7.79E-05
0.2	0.903885	0.903885	0.903885	0.903885	0.903885	0.904	1.30E-04
0.3	1.499234	1.499234	1.499234	1.499234	1.499234	1.499	1.71E-04
0.4	2.304105	2.304105	2.304105	2.304105	2.304105	2.304	1.90E-04
0.5	3.470132	3.470132	3.470132	3.470132	3.470132	3.470	2.25E-04
0.6	5.294825	5.294825	5.294825	5.294825	5.294825	5.294	2.43E-04
0.7	8.441305	8.441305	8.441305	8.441305	8.441305	8.442	3.05E-04
0.8	14.916937	14.916937	14.916937	14.916937	14.916937	14.917	3.22E-04
0.9	34.721476	34.721484	34.721484	34.721484	34.721484	34.722	5.17E-04
0.95	74.552341	74.619631	74.620917	74.620937	74.620937	74.621	7.11E-04

Table 3: Performance of Algorithm 2 in Heavy Traffic

Case		Minh and Sorli				
$\rho \backslash N$	1×10^4 2×10^4		$3 imes 10^4$	$4 imes 10^4$	$T = 1 \times 10^7$	
0.98	194.0544167173	194.5385548017	194.5559125683	194.5567071265	194.556	9.29E-04
	$5 imes 10^4$	$1 imes 10^5$	$2 imes 10^5$	$3 imes 10^5$		
0.98	194.5567179973	194.5567742874	194.5567742874	194.5567742874	194.556	9.29E-04
$\rho \setminus N$	$1 imes 10^4$	$3 imes 10^4$	$5 imes 10^4$	$1 imes 10^5$	T =	$1 imes 10^7$
0.99	372.0880005430	372.0880005430	391.8858614678	394.5238008176	394.532	1.45E-03
	$2 imes 10^5$	$3 imes 10^5$	$4 imes 10^5$	$5 imes 10^5$		
0.99	394.5331823499	394.5331886695	394.5331886695	394.5331886695	394.532	1.45E-03

In the setting of previous theorems,

 $W(D(1/p), RS(D(\rho), p)) \leq_{icx} W(D(1/p), M(\rho/p)),$ (2)

so that well known results for the D/M/1 queue yield

$$E[W(F_0, G_{u^*})] \le \frac{[2(1-\rho)\rho/(1-\delta)]c_a^2 + \rho^2 c_s^2}{2(1-\rho)},$$
(3)

where $\delta \in (0,1)$ solves the equation

$$\delta = \exp(-(1-\delta))/\rho). \tag{4}$$

 $E[W(F,G)] \leq E[W(F_0,G_{u^*})]$ (Tight UB)

$$\begin{array}{lcl} \mathcal{E}[\mathcal{W}(F,G)] &\leq & \mathcal{E}[\mathcal{W}(F_0,G_{u^*})](\mathsf{Tight~UB}) \\ &\leq & \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}(\mathsf{UB~Approx}) \end{array}$$

$$E[W(F,G)] \leq E[W(F_0, G_{u^*})] (\text{Tight UB}) \\ \leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)} (\text{UB Approx}) \\ < \frac{\rho^2([(2-\rho)c_a^2/\rho] + c_s^2)}{2(1-\rho)} (\text{Daley(1977)})$$

$$\begin{split} E[W(F,G)] &\leq E[W(F_0,G_{u^*})](\text{Tight UB}) \\ &\leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}(\text{UB Approx}) \\ &< \frac{\rho^2([(2-\rho)c_a^2/\rho]+c_s^2)}{2(1-\rho)}(\text{Daley(1977)}) \\ &< \frac{\rho^2([c_a^2/\rho^2]+c_s^2)}{2(1-\rho)}(\text{Kingman(1962)}) \end{split}$$

$$\begin{split} E[W(F,G)] &\leq E[W(F_0,G_{u^*})](\text{Tight UB}) \\ &\leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}(\text{UB Approx}) \\ &< \frac{\rho^2([(2-\rho)c_a^2/\rho]+c_s^2)}{2(1-\rho)}(\text{Daley(1977)}) \\ &< \frac{\rho^2([c_a^2/\rho^2]+c_s^2)}{2(1-\rho)}(\text{Kingman(1962)}) \end{split}$$

where $\delta \in (0,1)$ and $\delta = \exp(-(1-\delta)/\rho)$.

Table 4: A comparison of the bounds and approximations for the steady-state mean E[W] as a function of ρ for the case $c_a^2 = c_s^2 = 4.0$ and $c_s^2 = 4.0$.

ρ	Tight LB	HTA	Tight UB	UB Approx	δ	MRE	Daley	Kingman
0.10	0.00	0.044	0.422	0.422	0.000	0.003%	0.44	2.24
0.20	0.00	0.200	0.904	0.906	0.007	0.19%	1.00	2.60
0.30	0.00	0.514	1.499	1.51	0.041	0.60%	1.71	3.11
0.40	0.00	1.07	2.304	2.33	0.107	0.94%	2.67	3.87
0.50	0.25	2.00	3.470	3.51	0.203	1.15%	4.00	5.00
0.60	1.00	3.60	5.295	5.35	0.324	1.07%	6.00	6.80
0.70	2.42	6.53	8.441	8.52	0.467	0.93%	9.33	9.93
0.80	5.50	12.80	14.92	15.02	0.629	0.67%	16.00	16.40
0.90	15.25	32.40	34.72	34.84	0.807	0.35%	36.00	36.20
0.95	35.13	72.20	74.62	74.76	0.902	0.18%	76.00	76.10
0.98	95.05	192.1	194.6	194.7	0.960	0.07%	196.0	196.0
0.99	195.0	392.0	394.5	394.7	0.980	0.04%	396.0	396.0

Thank You!