Extremal Queueing Theory

Yan Chen

Advisor: Ward Whitt
Ph.D. Qualification Research Presentation
Department of Industrial Engineering and Operations Research

September 28, 2018

Objective: Develop a series of theoretical and numerical methodologies to explore "Extremal Queues Given Partial Information".

Important Idea: Focus on Queues instead of Formulas.

Objective: Develop a series of theoretical and numerical methodologies to explore "Extremal Queues Given Partial Information".

Important Idea: Focus on Queues instead of Formulas.
Objective: Develop a series of theoretical and numerical methodologies to explore "Extremal Queues Given Partial Information".
Important Idea: Focus on Queues instead of Formulas.

Objective: Develop a series of theoretical and numerical methodologies to explore "Extremal Queues Given Partial Information".

Important Idea: Focus on *Queues* instead of *Formulas*. 
The \textit{GI/GI/1} Model

We consider \textit{GI/GI/1} queue,

- a) \textbf{unlimited} waiting room and \textbf{FCFS} discipline;
- b) given mean and variance of $F$ and $G$;
- c) $F$ over bounded support $[0, M_a]$ and $G$ over $[0, M_s]$;
- d) traffic level < 1.

Let $W_n$ be $n$-th transient waiting time, assuming that the system starts empty with $W_0 = 0$. The sequence $\{W_n : n \geq 0\}$ is well known to satisfy the \textit{Lindley} recursion

$$W_{n+1} = [W_n + V_n - U_n]^+, \quad n \geq 0.$$  \hspace{1cm} (1)

Let $W = W_\infty$ denote the \textit{steady-state} waiting time.
We consider $GI/GI/1$ queue,

- a) unlimited waiting room and FCFS discipline;
- b) given mean and variance of $F$ and $G$;
- c) $F$ over bounded support $[0, M_a]$ and $G$ over $[0, M_s]$;
- d) traffic level $< 1$.

Let $W_n$ be $n$-th transient waiting time, assuming that the system starts empty with $W_0 = 0$. The sequence $\{W_n : n \geq 0\}$ is well known to satisfy the Lindley recursion

$$W_{n+1} = [W_n + V_n - U_n]^+, \quad n \geq 0.$$ (1)

Let $W = W_\infty$ denote the steady-state waiting time.
We consider $GI/GI/1$ queue,

- a) unlimited waiting room and FCFS discipline;
- b) given mean and variance of $F$ and $G$;
- c) $F$ over bounded support $[0, M_a]$ and $G$ over $[0, M_s]$;
- d) traffic level $< 1$.

Let $W_n$ be the $n$-th transient waiting time, assuming that the system starts empty with $W_0 = 0$. The sequence $\{W_n : n \geq 0\}$ is well known to satisfy the Lindley recursion

$$W_{n+1} = [W_n + V_n - U_n]^+, \quad n \geq 0.$$  

Let $W = W_\infty$ denote the steady-state waiting time.
The $GI/GI/1$ Model

We consider $GI/GI/1$ queue,

- a) unlimited waiting room and FCFS discipline;
- b) given mean and variance of $F$ and $G$;
- c) $F$ over bounded support $[0, M_a]$ and $G$ over $[0, M_s]$;
- d) traffic level $< 1$.

Let $W_n$ be $n$-th transient waiting time, assuming that the system starts empty with $W_0 = 0$. The sequence $\{W_n : n \geq 0\}$ is well known to satisfy the Lindley recursion

$$W_{n+1} = [W_n + V_n - U_n]^+, \quad n \geq 0.$$  \hspace{1cm} (1)

Let $W = W_\infty$ denote the steady-state waiting time.
The $GI/GI/1$ Model

We consider $GI/GI/1$ queue,

- a) unlimited waiting room and FCFS discipline;
- b) given mean and variance of $F$ and $G$;
- c) $F$ over bounded support $[0, M_a]$ and $G$ over $[0, M_s]$;
- d) traffic level $< 1$.

Let $W_n$ be $n$-th transient waiting time, assuming that the system starts empty with $W_0 = 0$. The sequence $\{W_n : n \geq 0\}$ is well known to satisfy the Lindley recursion

$$W_{n+1} = [W_n + V_n - U_n]^+, \quad n \geq 0. \quad (1)$$

Let $W = W_\infty$ denote the steady-state waiting time.
We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any $G$, what is the extremal inter-arrival time dist $F^*$ attaining the UB and LB of $E[W(F/G/1)]$?
- Given any $F$, what is the extremal service time dist $G^*$ attaining the UB and LB of $E[W(F/G/1)]$?
- What are the extremal $F^*$ and $G^*$ leading to overall UB and LB of $E[W(F/G/1)]$?
Classical $GI/GI/1$ Problem

We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any $G$, what is the extremal inter-arrival time dist $F^*$ attaining the UB and LB of $E[W(F/G/1)]$?
- Given any $F$, what is the extremal service time dist $G^*$ attaining the UB and LB of $E[W(F/G/1)]$?
- What are the extremal $F^*$ and $G^*$ leading to overall UB and LB of $E[W(F/G/1)]$?
We want to answer a **Long Standing Open Problem** in Queueing Theory.

- Given any $G$, what is the extremal inter-arrival time dist $F^*$ attaining the UB and LB of $\mathbb{E}[W(F/G/1)]$?
- Given any $F$, what is the extremal service time dist $G^*$ attaining the UB and LB of $\mathbb{E}[W(F/G/1)]$?
- What are the extremal $F^*$ and $G^*$ leading to overall UB and LB of $\mathbb{E}[W(F/G/1)]$?
Classical $GI/GI/1$ Problem

We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any $G$, what is the extremal inter-arrival time dist $F^*$ attaining the UB and LB of $\mathbb{E}[W(F/G/1)]$?
- Given any $F$, what is the extremal service time dist $G^*$ attaining the UB and LB of $\mathbb{E}[W(F/G/1)]$?
- What are the extremal $F^*$ and $G^*$ leading to overall UB and LB of $\mathbb{E}[W(F/G/1)]$?
Main Theoretical Results

Under the classical settings of $GI/GI/1$ model,

- Given any $G$ service time dist, $F^*(G)$ is a three-point dist.
- Given any $F$ inter-arrival time dist, $G^*(F)$ is a three-point dist.
- Under some regularity conditions, $F^*(G)$, $G^*(F)$ are two-point dists.

**Remark**

We also study the transient mean waiting time $\mathbb{E}[W_n]$ for $n \geq 1$ and unbounded support.
Main Theoretical Results

Under the classical settings of $GI/GI/1$ model,

- Given any $G$ service time dist, $F^*(G)$ is a three-point dist.
- Given any $F$ inter-arrival time dist, $G^*(F)$ is a three-point dist.
- Under some regularity conditions, $F^*(G), G^*(F)$ are two-point dists.

Remark

We also study the transient mean waiting time $\mathbb{E}[W_n]$ for $n \geq 1$ and unbounded support.
Main Theoretical Results

Under the classical settings of $GI/GI/1$ model,

- Given any $G$ service time dist, $F^*(G)$ is a three-point dist.
- Given any $F$ inter-arrival time dist, $G^*(F)$ is a three-point dist.
- Under some regularity conditions, $F^*(G), G^*(F)$ are two-point dists.

Remark

We also study the transient mean waiting time $\mathbb{E}[W_n]$ for $n \geq 1$ and unbounded support.
Main Theoretical Results

Under the classical settings of $GI/GI/1$ model,

- Given any $G$ service time dist, $F^*(G)$ is a three-point dist.
- Given any $F$ inter-arrival time dist, $G^*(F)$ is a three-point dist.
- Under some regularity conditions, $F^*(G)$, $G^*(F)$ are two-point dists.

**Remark**

We also study the transient mean waiting time $E[W_n]$ for $n \geq 1$ and unbounded support.
Main Theoretical Results

Under the classical settings of $GI/GI/1$ model,

- Given any $G$ service time dist, $F^*(G)$ is a three-point dist.
- Given any $F$ inter-arrival time dist, $G^*(F)$ is a three-point dist.
- Under some regularity conditions, $F^*(G)$, $G^*(F)$ are two-point dists.

Remark

*We also study the transient mean waiting time $E[W_n]$ for $n \geq 1$ and unbounded support.*
Overall Upper Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \leq \mathbb{E}[W(F_0/G_{u^*}/1)] \tag{2}
\]
\[
\leq \frac{2(1 - \rho)\rho/(1 - \delta)c_s^2 + \rho^2 c_s^2}{2(1 - \rho)} \tag{3}
\]
\[
< \text{Daley's Bound} < \text{Kingman’s Bound}. \tag{4}
\]

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

Overall Lower Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \geq \mathbb{E}[W(F_u/A_3/1)] \tag{5}
\]
\[
\geq \mathbb{E}[W(D/A_3/1)] = \text{Ott’s Bound}. \tag{6}
\]
Overall Upper Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \leq \mathbb{E}[W(F_0/G_u^*/1)] \\
\leq \frac{2(1 - \rho)\rho/(1 - \delta)c_a^2 + \rho^2c_s^2}{2(1 - \rho)} \\
< \text{Daley’s Bound} < \text{Kingman’s Bound.} \tag{4}
\]

\[\delta \in (0, 1) \text{ and } \delta = \exp(-(1 - \delta)/\rho).\]

Overall Lower Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \geq \mathbb{E}[W(F_u/A_3/1)] \geq \mathbb{E}[W(D/A_3/1)] = \text{Ott’s Bound.} \tag{6}
\]
Summary for Extremal $GI/GI/1$ Queues I

**Overall Upper Bound Inequalities:**

\[
\mathbb{E}[W(F/G/1)] \leq \mathbb{E}[W(F_0/G_{u*}/1)] \\
\leq \frac{2(1 - \rho)\rho/(1 - \delta)c_a^2 + \rho^2 c_s^2}{2(1 - \rho)} \\
< \text{Daley's Bound} < \text{Kingman's Bound.} \quad (4)
\]

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

**Overall Lower Bound Inequalities:**

\[
\mathbb{E}[W(F/G/1)] \geq \mathbb{E}[W(F_u/A_3/1)] \\
\geq \mathbb{E}[W(D/A_3/1)] = \text{Ott's Bound.} \quad (6)
\]
Overall Upper Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \leq \mathbb{E}[W(F_0/G_u^*/1)] 
\leq \frac{2(1 - \rho)\rho}{(1 - \delta)c^2_a + \rho^2 c^2_s} \frac{2(1 - \rho)}{2(1 - \rho)} 
< \text{Daley’s Bound} < \text{Kingman’s Bound.}
\]

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

Overall Lower Bound Inequalities:

\[
\mathbb{E}[W(F/G/1)] \geq \mathbb{E}[W(F_u/A_3/1)] 
\geq \mathbb{E}[W(D/A_3/1)] = \text{Ott’s Bound.}
\]
Summary for Extremal $GI/GI/1$ Queues II

Table: A comparison of the bounds and approximations for the scaled steady-state mean $(1 - \rho)E[W]/\rho^2$ in the $GI/GI/1$ model as a function of $\rho$ for the case $c_a^2 = c_s^2 = 4.0$.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Tight LB</th>
<th>HTA</th>
<th>Tight UB (2)</th>
<th>New UB (3)</th>
<th>$\delta$</th>
<th>MRE</th>
<th>Daley</th>
<th>Kingman</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.000</td>
<td>4.000</td>
<td>37.989</td>
<td>38.002</td>
<td>0.000</td>
<td>0.0%</td>
<td>40.000</td>
<td>202.000</td>
</tr>
<tr>
<td>0.20</td>
<td>0.000</td>
<td>4.000</td>
<td>18.080</td>
<td>18.112</td>
<td>0.007</td>
<td>0.2%</td>
<td>20.000</td>
<td>52.000</td>
</tr>
<tr>
<td>0.30</td>
<td>0.000</td>
<td>4.000</td>
<td>11.661</td>
<td>11.731</td>
<td>0.041</td>
<td>0.6%</td>
<td>13.333</td>
<td>24.222</td>
</tr>
<tr>
<td>0.40</td>
<td>0.000</td>
<td>4.000</td>
<td>8.641</td>
<td>8.722</td>
<td>0.107</td>
<td>0.9%</td>
<td>10.000</td>
<td>14.500</td>
</tr>
<tr>
<td>0.50</td>
<td>0.500</td>
<td>4.000</td>
<td>6.941</td>
<td>7.020</td>
<td>0.203</td>
<td>1.1%</td>
<td>8.000</td>
<td>10.000</td>
</tr>
<tr>
<td>0.60</td>
<td>1.111</td>
<td>4.000</td>
<td>5.884</td>
<td>5.946</td>
<td>0.324</td>
<td>1.1%</td>
<td>6.667</td>
<td>7.556</td>
</tr>
<tr>
<td>0.70</td>
<td>1.480</td>
<td>4.000</td>
<td>5.168</td>
<td>5.216</td>
<td>0.467</td>
<td>0.9%</td>
<td>5.714</td>
<td>6.082</td>
</tr>
<tr>
<td>0.80</td>
<td>1.719</td>
<td>4.000</td>
<td>4.662</td>
<td>4.693</td>
<td>0.629</td>
<td>0.7%</td>
<td>5.000</td>
<td>5.125</td>
</tr>
<tr>
<td>0.90</td>
<td>1.883</td>
<td>4.000</td>
<td>4.287</td>
<td>4.302</td>
<td>0.807</td>
<td>0.4%</td>
<td>4.444</td>
<td>4.469</td>
</tr>
<tr>
<td>0.95</td>
<td>1.946</td>
<td>4.000</td>
<td>4.134</td>
<td>4.142</td>
<td>0.902</td>
<td>0.2%</td>
<td>4.211</td>
<td>4.216</td>
</tr>
<tr>
<td>0.98</td>
<td>1.979</td>
<td>4.000</td>
<td>4.052</td>
<td>4.055</td>
<td>0.960</td>
<td>0.1%</td>
<td>4.082</td>
<td>4.082</td>
</tr>
<tr>
<td>0.99</td>
<td>1.990</td>
<td>4.000</td>
<td>4.025</td>
<td>4.027</td>
<td>0.980</td>
<td>0.0%</td>
<td>4.040</td>
<td>4.041</td>
</tr>
</tbody>
</table>
Summary for Effective Algorithms III

We develop effective **Numerical** and **Simulation** algorithms in Extremal Queues.

- **Queueing Reduction:**
  1) Deterministic Arrival Batches: $F_0/G_{u^*}/1$ to $D/RS(V,p)/1$.
  2) Daley's Decomposition for Service: $F_0/G_{u^*}/1$ to $F_0/D/1$.
- **Numerical Alg:** Negative Binomial (NB) & Discrete Time Markov Chain (DTMC).
- **Simulation Alg:** idle-time simulation algorithm.
We develop effective **Numerical** and **Simulation** algorithms in Extremal Queues.

- **Queueing Reduction:**
  1) **Deterministic Arrival Batches:** \( F_0/G_u^*/1 \) to \( D/RS(V,p)/1 \).
  2) **Daley’s Decomposition for Service:** \( F_0/G_u^*/1 \) to \( F_0/D/1 \).

- **Numerical Alg:** Negative Binomial (NB) & Discrete Time Markov Chain (DTMC).
- **Simulation Alg:** idle-time simulation algorithm.
We develop effective Numerical and Simulation algorithms in Extremal Queues.

- **Queueing Reduction:**
  1) **Deterministic Arrival Batches:** $F_0/G_u^*/1$ to $D/RS(V,p)/1$.
  2) **Daley’s Decomposition for Service:** $F_0/G_u^*/1$ to $F_0/D/1$.

- **Numerical Alg:** Negative Binomial (NB) & Discrete Time Markov Chain (DTMC).

- **Simulation Alg:** idle-time simulation algorithm.
We develop effective **Numerical** and **Simulation** algorithms in Extremal Queues.

- **Queueing Reduction:**
  1) *Deterministic Arrival Batches:* $F_0/G_u^*/1$ to $D/RS(V,p)/1$.
  2) *Daley’s Decomposition for Service:* $F_0/G_u^*/1$ to $F_0/D/1$.

- **Numerical Alg:** Negative Binomial (NB) & Discrete Time Markov Chain (DTMC).

- **Simulation Alg:** idle-time simulation algorithm.
The Possible Best Results

**Theorem**

(Counterexamples) Fix any service time dist $G$, $F^*(G) = F_0$; Fix any inter-arrival dist $F$, $G^*(F)$ is $G_0$ or $G_u$. The both arguments are **invalid**.

**Conjecture**

(Chen and Whitt I) Fix any $G$, the extremal $F^*(G)$ is a two-point distribution.

**Conjecture**

(Chen and Whitt II) Fix any two-point $F$, the extremal $G^*(F)$ is a two-point distribution.
The Possible Best Results

**Theorem**

(Counterexamples) Fix any service time dist $G$, $F^*(G) = F_0$; Fix any inter-arrival dist $F$, $G^*(F)$ is $G_0$ or $G_u$. The both arguments are invalid.

**Conjecture**

(Chen and Whitt I) Fix any $G$, the extremal $F^*(G)$ is a two-point distribution.

(Chen and Whitt II) Fix any two-point $F$, the extremal $G^*(F)$ is a two-point distribution.
The Possible Best Results

Theorem

(Counterexamples) Fix any service time dist $G$, $F^*(G) = F_0$; Fix any inter-arrival dist $F$, $G^*(F)$ is $G_0$ or $G_u$. The both arguments are invalid.

Conjecture

(Chen and Whitt I) Fix any $G$, the extremal $F^*(G)$ is a two-point distribution.

Conjecture

(Chen and Whitt II) Fix any two-point $F$, the extremal $G^*(F)$ is a two-point distribution.
Current Work

- Complete proof for conjectures and solve the whole problem.
- Study shape properties of $W$ in phase-type queues.
Current Work

- Complete proof for conjectures and solve the whole problem.
- Study shape properties of $W$ in phase-type queues.
Roadmap

- Establish Extremal Multi-server Queueing Theory
  long-standing open problem for $GI/GI/K$ for $K \geq 2$.
- Extension to Queues under Service Operations
  scheduling, staffing, readmission.
- Extension to Open and Closed Queueing Networks
  dependence of arrival and departure process information.
Establish Extremal Multi-server Queueing Theory
long-standing open problem for $GI/GI/K$ for $K \geq 2$.

Extension to Queues under Service Operations
scheduling, staffing, readmission.

Extension to Open and Closed Queueing Networks
dependence of arrival and departure process information.
Roadmap

- Establish Extremal Multi-server Queueing Theory
  long-standing open problem for $GI/GI/K$ for $K \geq 2$.
- Extension to Queues under Service Operations
  scheduling, staffing, readmission.
- Extension to Open and Closed Queueing Networks
  dependence of arrival and departure process information.
References

Kingman (1962)
Inequalities for the queue $GI/G/1$
_Biometrika._

Ward Whitt (1983)
Queueing Network Analyzer
_Bell System Technical Journal._

Ward Whitt (1984)
On Approximation for Queues, I: Extremal Distributions
_AT&T Bell Technical Journal._

Daley, Kreinin and Trengove (1992)
Inequalities concerning the waiting time in single-server queues: a survey
Queueing and Related Model, Clarendon Press.

Chen and Whitt (2018)
Extremal $GI/GI/1$ Queues Given First Two Moments
_Under Second Review of Operations Research._

Chen and Whitt (2018)
Algorithms for the Upper Bound Mean Waiting Time in the $GI/GI/1$ Queue
Submitted into INFORMS Journal of Computing.