Analysis of Hospital Networks via Time-varying Fluid Models with Blocking

Noa Zychlinski¹

Avishai Mandelbaum¹, Izack Cohen¹, Petar Momcilovic²

¹ The Faculty of Industrial Engineering and Management
Technion - Israel Institution of Technology

² Department of Industrial and Systems Engineering
University of Florida

May 2018
Motivation

• ‘Bed-Blocking’ Problem

Medical concern

Financial concern
NHS 'Bed Blocking' Reaches Three-Year High

Delays in patient assessments and allocating nursing home placements has led to increasing numbers of "bed blockers".

Hospital bed-blocking 'costs' NHS England £900m a year

About 8,500 “bed-blocking” patients are stuck in NHS hospitals every day, costing the health service £900m a year, a report says.

One in 20 patients ready to go – but stuck in hospital

'Bed-blocking' crisis putting strain on Scottish NHS

SOARING numbers of patients are waiting to be discharged from hospital beyond the official target, it emerged today as figures suggested the NHS is struggling.

Bed-blocking a massive problem

The crisis was raised by the Commons accounts committee, which said one in three patients was being delayed for more than a month before they were allowed to go home.
"300%" מהotros עוצמי בみたいな שיקומיות מהותית
למה Özma להדשיםلطיוול שיקומיים

ברזיל ישם רק 847 מירוצים שיקומיות, אך לעיתים ממה לכל אלף אדרחיה. דוב
שמאיור גאוןMALKH שמר רוכבי ייר מדרגות הולמה. והנה בן-7, שירדה
בה מקסיל לאפיזוד אדו מנייה לזרום לא הוארה, היא אוחדה עם מחוץ הואם במועד
שיטטורי להכנת במאז רקב זה לא מואד כלを持תי ביטויות השיקומי של

עמוסים רבים במבתי החולים. ב"ם
שיבץ בקוש לעלולぬ אל אלינו אתมงคลים

בבי"ם介质 השמר נרשמו תופסות של ייחר הת-30% בabee ו_minute
שעתים נUserController אתมงคลים במבתי חולים שארים. במחולקה להיניית של
היתשלומי 200% בכר נרשמו תפשות של כמות

משרד הבריאות: מספרımız של
הسيطرות יכלה בחור 25 שלנה

ה勐תום חישים שחשפים המשרד עלול כי מספר ביני ה-75 ביני ה-881
אלף ב-2039. ל kịchוז את, הפרסומת בסיוע של נשף יניקת ליא
的成长
Introduction

Life expectancy at birth

From: Israel Central Bureau of Statistics
"Life expectancy increases by five hours every day "
(Israel’s Ministry of Health, Jan 2017)
Introduction

Elderly percent from the total population 1960-2030

From: Israel Ministry of health, Information & Computing Department
Processed by: CBS
Introduction

1/5 of admissions to EDs are made by *elderly* patients (65+).

1/4 of elderly admissions to EDs are re-admissions.

1/3 of the hospitalizations are elderly patients.

<table>
<thead>
<tr>
<th>Life expectancy in nursing:</th>
<th>18 months</th>
<th>→</th>
<th>36 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>In 1995</td>
<td></td>
<td>→</td>
<td>in 2015</td>
</tr>
</tbody>
</table>
Patient Flow Chart

Hospital

Geriatric Institution

1.5 months
Skilled Nursing Care 4

5.5 months
Mechanical Ventilation 3

1 month
Geriatric Rehabilitation 2

Readmission
Death
Introduction

Average **waiting time** for a vacant bed

Source: our Partner hospital
Research Objectives

• Collecting and analyzing patient flow data between hospitals and geriatric institutions.

• Developing a mathematical model to describe the patient flow in the system.

• Offering ways to improve the performance of the system.
Literature Review

- High-level modeling of healthcare systems
- Queueing Networks with Blocking
- Bed planning for long-term care facilities
High-level modeling of healthcare systems

**Markov Models:** geriatric departments (Harrison and Millard, 1991), hospital and community (Taylor et al., 1997, 2000, Faddy and McClean, 2005, McClean and Millard, 2006), residential home care and nursing home care (Xie et al., 2005).

**System Dynamics:** analyze alternatives for shortening waiting times in UK NHS (Wolstenholme, 1999), analyze the dynamics of 'bed-blocking' in hospitals (Gray et al., 2006, Travers et al., 2008, Rohleder et al., 2013), forecast demand for geriatric services (Desai et al., 2008).

**Discrete Event Simulation:** analyze patient flow through geriatric departments (El-Darzi et al., 1998, Katsaliaki et al., 2005).
Literature Review

**Contribution:**

- Fluid model approach.
- Time-varying non-stationary analysis.

**High-level modeling of healthcare systems**

**Queueing Networks with Blocking**

**Bed planning for long-term care facilities**
Closed form expressions exist only for small, simple, single-server networks (Osorio and Bierlaire, 2009).

Approximations for complex networks:
The Decomposition Method (Takahashi et al., 1980, Koizumi et al., 2005, Osorio and Bierlaire, 2009)
The Expansion Method (Kerbache and MacGregor Smith, 1988, Cheah and Smith, 1994).

Literature Review

High-level modeling of healthcare systems

Queueing Networks with Blocking

Bed planning for long-term care facilities

Contribution:

- Fluid model approach.
- Time-varying non-stationary analysis.
- Modeling blocking without reflection.
- Analytical model for many-server networks.
- Time-varying analysis.
Literature Review

Bed planning for long-term care facilities

De Vries and Beekman, 1998 – a deterministic dynamic model for expressing the average waiting time for nursing homes.

Hare et al., 2009 – a deterministic model for predicting future long-term care needs.

Zhang et al., 2012 - heuristic algorithm integrated with discrete event simulation for capacity planning in nursing homes, in order to achieve target waiting time.

Ata et al., 2013 - a dynamic fluid model for profit maximization of hospice care under the Medicare reimbursement policy.

Jennings et al., 1997 – find the optimal number of private leased lines for profit maximization via a non-stationary analysis.
High-level modeling of healthcare systems

Queueing Networks with Blocking

Bed planning for long-term care facilities

**Contribution:**

- Fluid model approach.
- Time-varying non-stationary analysis.
- Modeling blocking *without* reflection.
- Analytic model for many-server networks.
- Time-varying analysis.
- Cost of bed-blocking.
- Mortality and readmission rates.
- Periodic reallocation.
Fluid Models

- Stochastic Discrete System

  In large overloaded systems

- Deterministic Continuous Model
Fluid Models
The Fluid Model

- Arrivals
- Inpatient ward
- Geriatric ward

<table>
<thead>
<tr>
<th>$\lambda(t)$</th>
<th>Arrival rate to Station 1 at time $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{12}$</td>
<td>Routing percentage from Station 1 to 2</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>Treatment rate at Station $i$</td>
</tr>
<tr>
<td>$N_i$</td>
<td>Number of beds at Station $i$</td>
</tr>
</tbody>
</table>
The Fluid Model

\[ q_i(t) \] Number of patients at Station \( i \)

\[ x_1(t) \] Number of arrivals to Station 1 that have not completed their treatment at Station 1 at time \( t \)

\[ x_2(t) \] Number of patients that have completed treatment at Station 1, but not at Station 2 at time \( t \)
The Fluid Model

\[ q_1(t) = x_1(t) + \left( x_2(t) - N_2 \right)^+; \]

\( (x)^+ = \max(x, 0) \)
The Fluid Model

\[ q_1(t) \]

\[ q_2(t) \]

\[ x_1(t) \]

\[ x_2(t) \]

\[ q_2(t) = x_2(t) \wedge N_2 \]

\[ x \wedge y = \min(x, y) \]
The Fluid Equations

\[ \dot{x}_1(t) = \text{Arrival rate} - \text{Departure rate} \]

\[ \dot{x}_1(t) = \lambda(t) - \mu_1 \left[ \text{Occupied unblocked beds in Station 1} \right] \]
The Fluid Equations

$\dot{x}_1(t) = \text{Arrival rate} - \text{Departure rate}$

$\dot{x}_1(t) = \lambda(t) - \mu_1 \left[ x_1(t) \wedge \left( \text{Unblocked beds in Station 1} \right) \right]$
The Fluid Equations

$x_1(t) = \text{Arrival rate} - \text{Departure rate}$

$x_1(t) = \lambda(t) - \mu_1 \left[ x_1(t) \land \left( N_1 - \right) \right]$  \hspace{1cm} \text{Blocked beds}

Unblocked beds in Station 1

Occupied unblocked beds in Station 1
The Fluid Equations

\[ \dot{x}_1(t) = \text{Arrival rate} - \text{Departure rate} \]

\[ \dot{x}_1(t) = \lambda(t) - \mu_1 \left[ x_1(t) \land \left( N_1 - (x_2(t) - N_2)^+ \right) \right] \]

- **Blocked beds**
- **Unblocked beds in Station 1**
- **Occupied unblocked beds in Station 1**
The Fluid Equations

\[ \dot{x}_2(t) = p_{12} \cdot \mu_1 \left[ x_1(t) \wedge \left( N_1 - (x_2(t) - N_2)^+ \right) \right] - \mu_2 (x_2(t) \wedge N_2) \]

Arrival rate

Departure rate
Network Modeling

<table>
<thead>
<tr>
<th>$p_{1i}$</th>
<th>Routing percentage from Station 1 to Station $i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_i$</td>
<td>Readmission rate from Station $i$</td>
</tr>
<tr>
<td>$\theta_i$</td>
<td>Mortality rate at Station $i$</td>
</tr>
</tbody>
</table>
Network Modeling

Exogenous arrival rate

Inpatient Wards 1

Patient Condition

Skilled Nursing Care 4
Mechanical Ventilation 3
Geriatric Rehabilitation 2

\[
\dot{x}_1(t) = \lambda(t) + \sum_{i=2}^{4} \beta_i(x_i(t) \land N_i) - \theta_1 x_1(t) - \mu_1 \left[ x_1(t) \land \left( N_1 - \sum_{i=2}^{4} (x_i(t) - N_i)^+ \right) \right]
\]

Internal arrival rate
Network Modeling

Hospital

Geriatric Institution

\[ \dot{x}_1(t) = \lambda(t) + \sum_{i=2}^{4} \beta_i(x_i(t) \wedge N_i) - \theta_1 x_1(t) - \mu_1 \left[ x_1(t) \wedge \left( N_1 - \sum_{i=2}^{4} (x_i(t) - N_i)^+ \right) \right] \]

Mortality rate

Treatment completion rate
\[ \dot{x}_i(t) = p_{1i} \cdot \mu_1 \left[ x_1(t) \wedge \left( N_1 - \sum_{i=2}^{4} (x_i(t) - N_i)^+ \right) \right] - \beta_i (x_i(t) \wedge N_i) - \theta_i x_i(t) - \mu_i (x_i(t) \wedge N_i), \]

\( i = 2, 3, 4. \)
1. Patient flow data from a district comprising 4 hospitals and 3 geriatric institutions (Israel’s Ministry of Health).

2. Waiting lists for geriatric wards, including individual waiting times (our Partner Hospital).
Waiting-List Length in Hospital Model vs. Data

\[
\begin{align*}
\mu_1 &= 1/4.85, \quad \mu_2 = 1/30, \quad \mu_3 = 1/160, \quad \mu_4 = 1/45, \\
\beta_2 &= 1/250, \quad \beta_3 = 1/1000, \quad \beta_4 = 1/1000, \\
\theta_1 &= 1/125, \quad \theta_2 = 1/2500, \quad \theta_3 = 1/1000, \quad \theta_4 = 1/1000, \\
N_1 &= 600, \quad N_2 = 234, \quad N_3 = 93, \quad N_4 = 120.
\end{align*}
\]
Estimating the Optimal Number of Geriatric Beds

**Objective:**
To minimize the total cost of operating the system

\[ T \] – Planning horizon (usually 3-5 years)

**Decision variables:**
\[ N_i \] – Number of beds at Ward \( i, \ i = 2, 3, 4 \)
Estimating the Optimal Number of Geriatric Beds

Cost construction

\[ C_o \]
per bed per day

Overage cost

\[ C_u \]
per bed per day

Underage cost
Estimating the Optimal Number of Geriatric Beds

The objective function:

$$C^{(0)}(N_2, N_3, N_4) = \sum_{i=2}^{4} \int_0^T \left[ C_{ui} \cdot b_i(t) + C_{oi} \cdot (N_i - q_i(t))^+ \right] dt$$

Blocked patients    Empty beds

Analytically intractable
The Offered-Load – \( r(t) \)

The average amount of work being processed at each station, under the assumption of an infinite number of servers (beds).

In our case, we have to consider: treatment, mortality and readmission rates.

The calculation is done by solving the fluid equations with \( N_i \equiv \infty, \ i = 2, 3, 4 \)

\[
\dot{r}_1(t) = \lambda(t) + \sum_{i=2}^{4} \beta_i r_i(t) - \theta_1 r_1(t) - \mu_1 (r_1(t) \wedge N_1)
\]

\[
\dot{r}_i(t) = p_{1i}(t) \cdot \mu_1 (r_1(t) \wedge N_1) - (\beta_i + \theta_i + \mu_i) r_i(t), \quad i = 2, 3, 4.
\]
Estimating the Optimal Number of Beds

\[ C(N) = \int_0^T \left[ C_u \cdot (r(t) - N)^+ + C_o \cdot (N - r(t))^+ \right] dt \]
We use the approach of Jennings et al. (1997) and treat $N$ as a continuous variable. We let

$$r_d = \{r_d(t) \mid 0 \leq t \leq T\}$$

denote the decreasing rearrangement of $r$ on the interval $[0, T]$. 

Estimating the Optimal Number of Beds
Theorem:

The optimal number of beds, which minimizes $C(N)$ is:

$$N^* = r_d \left( \frac{C_o}{C_o + C_u} T \right)$$

When $C_o = 0$ and $C_u = 0$.
Optimal Number of Beds

\[ C_{u2} = 2.667 \ C_{o2}, \quad C_{u3} = 1.882 \ C_{o3}, \quad C_{u4} = 4.267 \ C_{o4} \]
Optimal Number of Beds

\[ C_{u2} = 2.667 \, C_{o2}, \quad C_{u3} = 1.882 \, C_{o3}, \quad C_{u4} = 4.267 \, C_{o4} \]
Optimal Number of Beds

Increasing the current number of beds by

25%, 35% and 33% in Rehabilitation, Mechanical Ventilation and Skilled Nursing Care

Overage and underage cost reduction of

51%, 53% and 69%

Waiting list length reduction of

67%, 74% and 88%
Waiting List Length in Hospital

Optimal Solution

![Diagram showing waiting list length over time for different services: Rehabilitation, Mechanical Ventilation, and Skilled Nursing. The graph has x-axis labeled 't [days]' and y-axis labeled 'Waiting list length.' The peaks indicate varying levels of demand throughout the first three years.]
# Comparing Optimal Solutions

<table>
<thead>
<tr>
<th>Ward</th>
<th>Fluid Model</th>
<th>Offered-load</th>
<th>Simulation</th>
<th>Maximal difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C^{(0)}(N_2, N_3, N_4)$</td>
<td>$C(N_2, N_3, N_4)$</td>
<td>Simulation</td>
<td>$N^*$ (Total cost)</td>
</tr>
<tr>
<td></td>
<td>$N^*$ (Total cost)</td>
<td>$N^*$ (Total cost)</td>
<td>$N^*$ (Total cost)</td>
<td>$N^*$ (Total cost)</td>
</tr>
<tr>
<td>Rehabilitation</td>
<td>295 (2,601,667)</td>
<td>292 (2,683,042)</td>
<td>294 (2,633,167)</td>
<td>1.0% (3.0%)</td>
</tr>
<tr>
<td>Mechanical Ventilation</td>
<td>128 (1,493,917)</td>
<td>126 (1,547,000)</td>
<td>128 (1,499,167)</td>
<td>1.6% (3.4%)</td>
</tr>
<tr>
<td>Skilled Nursing</td>
<td>161 (1,213,333)</td>
<td>159 (1,226,750)</td>
<td>160 (1,215,667)</td>
<td>1.3% (1.1%)</td>
</tr>
<tr>
<td>Total Number of beds</td>
<td>584 (5,308,917)</td>
<td>577 (5,456,792)</td>
<td>582 (5,348,000)</td>
<td>1.2% (2.7%)</td>
</tr>
</tbody>
</table>
Optimal Number of Beds

![Graph showing the optimal number of beds with overage periods marked.](image)

- Rehabilitation
- Mechanical Ventilation
- Skilled Nursing

R(t) / N

1st year
2nd year
3rd year

t [days]
Periodic Reallocation of Beds

Underage cost

Overage cost
Optimal Periodic Reallocation

The decisions:

1.) The number of periods each year (reallocation points).

2.) The length of each period.

3.) The number of beds in each period.

$C_R$ - Reallocation cost associated with adding and removing a bed.
Optimal Periodic Reallocation

Not Including Reallocation Costs
Optimal Periodic Reallocation

Including Reallocation Costs

![Graph showing the trend of r(t) over the first, second, and third years with dashed lines representing Rehabilitation, Mechanical Ventilation, and Skilled Nursing.](image-url)
Optimal Periodic Reallocation

Quarterly Reallocation

Graph showing the periodic reallocation over time with distinct lines for Rehabilitation, Mechanical Ventilation, and Skilled Nursing.
Waiting List Length in Hospital

Quarterly Reallocation

![Graph showing waiting list length over time for different categories: Rehabilitation, Mechanical Ventilation, and Skilled Nursing.](image)
The Bed Blocking Problem

Blocking

ED

Hospital wards

Geriatric wards

\[ \mu_1, N_1 \]

\[ \mu_2, N_2 \]

\[ \mu_{k-1}, N_{k-1} \]

\[ \mu_k, N_k \]

\[ p_{12} \]

\[ p_{23} \]

\[ p_{k-1,k} \]

\[ 1-p_{12} \]

\[ 1-p_{23} \]

\[ 1-p_{k-1,k} \]
K Stations in Tandem

The general unified model
Analysis Time-varying Tandem Queues with Blocking:
Modeling, Analysis and Operational Insights via Fluid Models with Reflection
Two Stations with No Waiting Rooms

The Stochastic Processes:

\[ X = \{X_1(t), X_2(t), t \geq 0\} \]

\[ Q = \{Q_1(t), Q_2(t), t \geq 0\} \]
The Stochastic Model

$X_1(t) = X_1(0)$

Arrivals

Transfers to Station 2

Departures from Station 1

Blocked customers who were forced to leave the system
The Stochastic Model

\[ X_1(t) = X_1(0) \]

Arrivals

Transfers to Station 2

Departures from Station 1

 Blocked customers who were forced to leave the system
The Stochastic Model

Station 1 is full

Blocked customers who were forced to leave the system

Arrival process

$X_1(t) = X_1(0) + A(t) - D_1 \left( p \mu_1 \int_0^t X_1(u) \, du \right) - D_3 \left( (1 - p) \mu_1 \int_0^t X_1(u) \, du \right)$

$\int_0^t 1 \left\{ x_1(u-) + (x_2(u-) - N_2)^+ = N_1 \right\} \, dA(u)$

$L(t)$

$D_i(t) - Standard Poisson process$
The Stochastic Model

\[ X_2(t) = X_2(0) + D_1 \left( p \mu_1 \int_0^t X_1(u) du \right) - D_2 \left( \mu_2 \int_0^t [X_2(u) \wedge N_2] du \right) \]
The Reflection Representation

\[ \begin{align*}
&\begin{cases}
X_1(t) \\
X_1(t) + X_2(t)
\end{cases}
= \begin{cases}
Y_1(t) - L(t) \\
Y_2(t) - L(t)
\end{cases} \\
&\leq \begin{cases}
\frac{N_1}{N_1 + N_2}
\end{cases}, \quad t \geq 0,
\end{align*} \]

\[ \int_0^\infty \frac{1}{\{X_1(t) + (X_2(t) - N_2)^+ < N_1\}} \, dL(t) = 0, \]

when Station 1 is full
Geometric Representation of the Reflection

\[ X_1 + X_2 = N_1 + N_2 \]

\[ X_1 = N_1 \]
The Reflection Representation - R

\[ R_1(t) = N_1 - X_1(t) \text{, Idle or blocked servers in Station 1,} \]
\[ R_2(t) = N_1 + N_2 - (X_1(t) + X_2(t)) \text{, Idle servers in Stations 1 and 2} \]

\[
\begin{aligned}
\begin{bmatrix}
R_1(t) \\
R_2(t)
\end{bmatrix}
&= 
\begin{bmatrix}
\tilde{Y}_1(t) + L(t) \\
\tilde{Y}_2(t) + L(t)
\end{bmatrix} \\
&\geq 0, \quad t \geq 0, \\
\end{aligned}
\]

\[
\begin{aligned}
&dL \geq 0, \quad L(0) = 0, \\
&\int_0^\infty 1_{\{R_1(t) \land R_2(t) > 0\}} \, dL(t) = 0,
\end{aligned}
\]

When \( R_1(t) = 0 \) or \( R_2(t) = 0 \)

\[ L(t) \]
Equal positive displacements in both $R_1$ and $R_2$
Representation in Terms of Reflection

\[
\begin{cases}
\begin{pmatrix}
R_1(t) \\
R_2(t)
\end{pmatrix} = \begin{pmatrix}
\tilde{Y}_1(t) + L(t) \\
\tilde{Y}_2(t) + L(t)
\end{pmatrix} \geq 0, & t \geq 0, \\
dL \geq 0, & L(0) = 0, \\
\int_0^\infty 1 \left\{ R_1(t) \wedge R_2(t) > 0 \right\} \, dL(t) = 0,
\end{cases}
\]

\[L(t) = \sup_{0 \leq s \leq t} \left( - (\tilde{Y}_1(s) \wedge \tilde{Y}_2(s)) \right)^+\]
The Fluid Limit – Two Stations

\[ x_1(t) = x_1(0) - \mu_1 \int_0^t \left[ x_1(u) \wedge (N_1 - b(u)) \right] du \]

\[ + \int_0^t \left[ 1\{x_1(u) < N_1 + H\} \cdot 1\{x_1(u) + x_2(u) < N_1 + N_2 + H\} \cdot \lambda(u) \right] du \]

\[ + \int_0^t \left[ 1\{x_1(u) = N_1 + H\} \cdot 1\{x_1(u) + x_2(u) < N_1 + N_2 + H\} \cdot \lambda(u) \wedge l_1^*(u) \right] du \]

\[ + \int_0^t \left[ 1\{x_1(u) < N_1 + H\} \cdot 1\{x_1(u) + x_2(u) = N_1 + N_2 + H\} \cdot \lambda(u) \wedge l_2^*(u) \right] du \]

\[ + \int_0^t \left[ 1\{x_1(u) = N_1 + H\} \cdot 1\{x_1(u) + x_2(u) = N_1 + N_2 + H\} \cdot \lambda(u) \wedge l_1^*(u) \wedge l_2^*(u) \right] du, \]

\[ x_2(t) = x_2(0) + \int_0^t \left[p\mu_1 [x_1(u) \wedge (N_1 - b(u))] - \mu_2 (x_2(u) \wedge N_2) \right] du, \]

\[ l_1^*(u) = \mu_1 N_1, \]

\[ l_2^*(u) = \mu_2 N_2 + (1 - p)\mu_1 \left(x_1(u) \wedge (N_1 - b(u))\right), \]

\[ b(u) = (x_2(u) - N_2)^+. \]

\[ q_1(t) = x_1(t) + b(t); \]

\[ q_2(t) = x_2 \wedge N_2. \]
Total Number in each Station - A Numerical Example

\[ N_1 = 200, \; N_2 = 150, \; \mu_1 = \frac{1}{10}, \; \mu_2 = \frac{1}{20}, \; p = 1, \; q_1(0) = q_2(0) = 0. \; \text{and} \; \lambda(t) = 2t, \; 0 \leq t \leq 120. \]
The $G_t/M/(N+H)$ Queue

$$q(t) = q(0) + \int_0^t \left[ \lambda(u) \cdot 1_{\{q(u)<N+H\}} + [\lambda(u) \wedge \mu N] \cdot 1_{\{q(u)=N+H\}} - \mu [q(u) \wedge N] \right] du$$

K Stations in Tandem
Numerical Experiments and Operational Insights
Line-Length Effect – Infinite Capacities

\[ \lambda(t) = \bar{\lambda} + \beta \sin(\gamma t), \quad t \geq 0 \text{ with } \bar{\lambda} = 9, \beta = 8 \text{ and } \gamma = 0.02, \mu_i = 1/20 \text{ and } q_i(0) = 0, \ i = 1, \ldots, 8. \]

Line-Length Effect – Finite Capacities

Identical stations

\[
\lambda(t) = \bar{\lambda} + \beta \sin(\gamma t), \quad t \geq 0, \quad \bar{\lambda} = 9, \quad \beta = 8 \quad \text{and} \quad \gamma = 0.02, \quad N_i = 200, \quad \mu_i = 1/20 \quad H = \infty
\]
Numerical Experiments

Case 1: No waiting room \((H = 0)\)

Case 2: An infinite sized waiting room \((H = \infty)\)
Total Number of Customers in Each Station

k=8 identical stations

Case 1
No waiting room
\( H = 0 \)

Average sojourn time is 20% shorter

Case 2
An infinite sized waiting room
\( H = \infty \)

\[ N = 200, \mu = \frac{1}{20}, p_{i,i+1} = 1, q_i(0) = 0, i = 1, 2, \ldots, 8 \] and \( \lambda(t) = 2t, \quad 0 \leq t \leq 40. \]
The Bottleneck Effect on Performance

Bottleneck = Station 8

**Case 1**
No waiting room
\[ H = 0 \]

**Case 2**
An infinite sized waiting room
\[ H = \infty \]

\[ N_i = 200, \mu_i = 1/20, i = 1, 2, \ldots, 7, \quad N_8 = 120, \mu_8 = 1/40 \quad \text{and} \quad \lambda(t) = 2t, \quad 0 \leq t \leq 40. \]
The Bottleneck Effect on Performance

Bottleneck = Station 8

Case 1
No waiting room
\( H = 0 \)

Case 2
An infinite sized waiting room
\( H = \infty \)

\( N_i = 200, \mu_i = 1/20, \ i = 1, 2, \ldots, 7, \ N_8 = 120, \mu_8 = 1/40 \) and \( \lambda(t) = 2t, \ 0 \leq t \leq 40. \)
Bottleneck Location Effect on Performance

\[ H = \infty \]

**Bottleneck = Station 1**

**Bottleneck = Station 8**

\[ N_i = 200, \ \mu_i = 1/20, \ i = 1, 2, \ldots, 7, \ \ N_8 = 120, \ \mu_8 = 1/40 \ \text{and} \ \lambda(t) = 2t, \ 0 \leq t \leq 40. \]
Q: In which setting will the sojourn time be shorter?
Sojourn Time - 8 Stations

\[ N_i = 200, \mu_i = 1/20, i = 1, 2, \ldots, 8, i \neq j, N_j = 120, \mu_j = 1/40. \lambda(t) = 20, 0 \leq t \leq 100 \]
Sojourn Time

Waiting time before Station $1$ + Blocking time at Stations $1, \ldots, k-1$ + Service time at Stations $1, \ldots, k$
Bottleneck Location and Waiting Room Size

8 Stations

Blocking Time

Waiting Time

\[ N_i = 200, \mu_i = 1/20, i = 1, 2, \ldots, 8, i \neq j, \quad N_j = 120, \mu_j = 1/40. \lambda(t) = 20, 0 \leq t \leq 100 \]
Blocking Time in each Station (H=0)

8 Stations

$N_i = 200, \mu_i = 1/20, i = 1, 2, \ldots, 8, i \neq j, N_j = 120, \mu_j = 1/40. \lambda(t) = 20, 0 \leq t \leq 100$
Blocking Time in each Station (H=0)

8 Stations

\[
N_i = 200, \quad \mu_i = 1/20, \quad i = 1, 2, \ldots, 8, \quad i \neq j, \quad N_j = 120, \quad \mu_j = 1/40. \quad \lambda(t) = 20, \quad 0 \leq t \leq 100
\]
Blocking Time in each Station (H=0)

8 Stations

\[ N_i = 200, \mu_i = 1/20, i = 1, 2, \ldots, 8, \; i \neq j, \; N_j = 120, \mu_j = 1/40. \lambda(t) = 20, \; 0 \leq t \leq 100 \]
Analysis Time-Varying Many-Server Finite-Queues in Tandem: Comparing Blocking Mechanisms via Fluid Models
Motivation

Surgery rooms

Recovery rooms
Blocking Mechanisms

**Blocking After Service (BAS)**

![Blocking After Service Diagram]

**Blocking Before Service (BBS)**

![Blocking Before Service Diagram]
Blocking Mechanisms

Blocking After Service (BAS)

Blocking Before Service (BBS)
Blocking Mechanisms

BAS

BBS
Blocking Before Service (BBS)

\[ Q_1(t) = Q_1(0) + A(t) - \int_0^t 1\{Q_1(u-) = H_1 + N_1\} dA(u) \]

\[ - D_1 \left( \mu_1 \int_0^t [Q_1(u) \wedge N_1 \wedge (H_2 + N_2 - Q_2(u))] du \right) \]

Available capacity at Station 2
Customers in service at Station 1

\[ D_i(t) - \text{Standard Poisson process} \]
Line-Length Effect – Finite Capacities

Blocking After Service (BAS)  Blocking Before Service (BBS)

\[ \lambda(t) = 9 + 8 \sin(0.02t), \quad H_1 = \infty, \quad H_j = 0, \ j=2,\ldots,k, \quad N_i = 200, \mu_i = 1/20, \ i=1,\ldots,k. \]
Line-Length Effect – Finite Capacities

Blocking **After Service (BAS)**

Blocking **Before Service (BBS)**

\[ \lambda(t) = 9 + 8 \sin(0.02t), \quad H_1 = \infty, \quad H_j = 200, \quad j=2,..,k, \quad N_i = 200, \quad \mu_i = \frac{1}{20}, \quad i=1,..,k. \]
Steady-State Performance

\[ \lambda(t) \equiv \lambda, \ t \geq 0, \ q_i(0) = q_i(t) \equiv \bar{q}_i, \ \forall t \geq 0, \ i = 1, \ldots, k. \]

**Blocking After Service (BAS)**

\[ \delta_{\text{BAS}} = \left( \lambda \bigwedge_{j=1}^{k} \mu_j N_j \right) \]
- Arrival rate
- Processing capacity of the bottleneck

**Blocking Before Service (BBS)**

\[ \delta_{\text{BBS}} = \left( \lambda \bigwedge_{j=1}^{k} \mu_j N_j \right) \bigwedge_{j=2}^{k} \frac{H_j + N_j}{1/\mu_{j-1} + 1/\mu_j} \]
- Arrival rate
- Processing capacity of the bottleneck
- Processing capacity of the "virtual" bottleneck
Steady-State Performance

\[ \delta_{BBS} \leq \delta_{BAS} \]

\[ \delta_{BBS} = \delta_{BAS} \text{ when} \]

\[ \lambda \land \bigwedge_{j=1}^{k} \mu_j N_j \leq \bigwedge_{j=2}^{k} \frac{H_j + N_j}{1/\mu_{j-1} + 1/\mu_j} \]
### Numerical Example

**Scenario 1**

<table>
<thead>
<tr>
<th></th>
<th>Surgery rooms</th>
<th>Recovery rooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service rate (μ)</td>
<td>1/60</td>
<td>1/60</td>
</tr>
<tr>
<td>Number of beds (N)</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Waiting room size (H)</td>
<td>[0,∞)</td>
<td>0</td>
</tr>
</tbody>
</table>

δ_{BBS} = δ_{BAS}  
10 customers per hour
## Numerical Example

### Scenario 2

<table>
<thead>
<tr>
<th></th>
<th>Surgery rooms</th>
<th>Recovery rooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service rate ((\mu))</td>
<td>1/60</td>
<td>1/120</td>
</tr>
<tr>
<td>Number of beds (N)</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Waiting room size (H)</td>
<td>([0, \infty))</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \delta_{\text{BBS}} < \delta_{\text{BAS}} \]

- 6.67 per hour
- 10 per hour
Average Sojourn Time

Calculate the number of customers in each station:

\[ i = \min \left\{ \arg \min_{j=1}^{k} \mu_j N_j, \arg \min_{j=2}^{k} \frac{H_j + N_j}{1/\mu_{j-1} + 1/\mu_j} \right\} \]

When \( \delta = \lambda \), then \( \bar{q}_j = \lambda/\mu_j \), \( j = 1, \ldots, k \). Otherwise (when \( \delta < \lambda \)),

- \( \bar{q}_1 = H_1 + N_1; \)
- \( \bar{q}_j = H_j + N_j - \delta/\mu_{i-1} \), \( j = 2, \ldots, i; \)
- \( \bar{q}_j = \delta/\mu_j \), \( j = i + 1, \ldots, k \),

Calculate the average sojourn time by Using Little’s Law

\[
\text{Throughput} = \frac{\text{Total number of customers}}{\text{Throughput}}
\]

The first actual/virtual bottleneck
Numerical Example

Scenario 2

<table>
<thead>
<tr>
<th></th>
<th>Surgery rooms</th>
<th>Recovery rooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service rate (µ)</td>
<td>1/60</td>
<td>1/120</td>
</tr>
<tr>
<td>Number of beds (N)</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Waiting room size (H)</td>
<td>[0,∞)</td>
<td>0</td>
</tr>
</tbody>
</table>

Average sojourn time:
- δBBS: 6.67 per hour
- δBAS: 10 per hour

Average sojourn time:
- 3.5 hours
- 3 hours
Network Throughput in Steady-State

Blocking **After** Service (BAS)

\[ \delta^{BAS} = \lambda \land \bigwedge_{j=1}^{k} \mu_j N_j \]

Blocking **Before** Service (BBS)

\[ \delta^{BBS} = \lambda \land \bigwedge_{j=1}^{k} \mu_j N_j \land \bigwedge_{j=2}^{k} \frac{H_j + N_j}{1/\mu_{j-1} + 1/\mu_j} \]
Output Rate

\( H_1 = 0 \)

\[ \lambda(t) = 2t, \ t \leq 40, \ N_i = 200, \mu_i = 1/20, \ i=1,...,k. \]
Future Directions:

- **Exploratory Data Analysis (EDA)**

Conducting an integrative analysis:

- Relating transfer delays to readmission rates, treatment durations and patient clinical condition.
- Relating blocking of geriatric wards to ED boarding time.
Future Directions:

- Analyzing the option of adding an intermediate, sub-acute geriatric ward (a step down unit).
THANK YOU!

noazy@tx.technion.ac.il