On Measuring, Modeling and Analyzing Healthcare Systems (Hospitals, . . . , Call Centers) in Real-Time

From Small Measurements through Big Data to Analytics

Avishai Mandelbaum

IE&M & SEELab, Technion

- **Lecture will be downloadable from my Technion website**

http://ie.technion.ac.il/serveng

Research Partners

- **Students**:
  Aldor*, Baron Yonit*, Carmeli Boaz*, Chen Hong*, Cohen Izik*, Feldman Zohar*, Garnett*, Ghebali*, Gurvich*, Khudiakov*, Koren, Maman*, Marmor*, Reich*, Rosenshmidt*, Shaikhet*, Senderovic, Tseytlin*, Yom-Tov*, Carmeli-Yuviler, Zaied*, Zeltyn*, Zychlinski, Zohar Eti*, Zviran*
- **Theory**:
  Atar, Jelenkovic, Massey, Reiman, Shimkin, Stolyar, Whitt, Zhang Hanqin[1]
- **Exploratory Data Analysis, Data Sources**:
  He Shuangchi, Kim Song-Hee, Liberman, Liu Nan, Sim Melvyn, Tezcan, Won Chul Cha, Wu Dan, Ye Han, . . .
- **Data-based Theory**:
  Armony, Azriel, Brown, Carmeil-Yuviler, Cohen Izik, Feigin, Goldberg, Gal, Gans, Gorfine, Gurvich, Huang Junfei, Jansen, Kaspi, van Leeuwaarden, Marmor, Mathisjen, Momcilovic, Parush, Ritov, Senderovic, Shen Haipeng, Tseytlin, Wasserkrug, Weidlich, Yom-Tov, Zeltyn, Zhao Linda, Zhang Jiheng, Zwart, . . .
- **Industry**:
  Mizrahi Bank, Fleet Bank, Rambam Hospital, IBM Research, Hapoalim Bank, Pelephone Cellular, Samsung Hospital, Alexandra Health, DayHospital . . .
- **Technion SEE Laboratory (SEELab)**:
  Feigin; Trofimov, Nadjharov, Gavako; Lab Alumni; Research Assistants

[1] In many Western countries, there is a short list of popular “first names," but countless “last names." In China, it is just the reverse. The list of last names is short, and the number of first names is in the billions (from chinapage.com/biography/lastname.html).

Research Goals

- Reality: Service Systems
  e.g. **Hospitals, Call Centers**, Websites, . . .

- Models = **ServNets**

  Simple models at the service of complex realities

  Q-Nets, F-Nets, D-Nets, Sim-Nets: Operations Research (Math)

  but also **P-Nets, C-Nets**, **. . .**, as in BPM, WFM, Process Mining

Research Goals

- Reality: Service Systems
  e.g. **Hospitals, Call Centers**, Websites, . . .

- Models = **ServNets**

  Simple models at the service of complex realities

  Q-Nets, F-Nets, D-Nets, Sim-Nets: Operations Research (Math)

  but also **P-Nets, C-Nets**, **. . .**, as in BPM, WFM, Process Mining

- Goals:

  Data-based creation, validation, analysis & control

  of ServNets,

  on demand & automatically

Feasible? Why be Optimistic?

e.g. Emergency-Department simulation, queueing model & analysis of a specific ED:

- Online: create **model(s)** of ED on normal Monday mornings
- Real-time: **control** patient-flow bottlenecks
- Short-term: on Monday, set Tuesday's **staffing** levels (or next week's)
- **Design**: physician protocol – explore (triage) vs. exploit (discharge)

Technion IE&M research (OR/QS/IE) + **SEELab** technologies:

- **Operational recipes** (control, staffing, design) via Math-Nets (Q, F, D)

Feasible? Why be Optimistic?

e.g. Emergency-Department simulation, queueing model & analysis of a specific ED:

- Online: create **model(s)** of ED on normal Monday mornings
- Real-time: **control** patient-flow bottlenecks
- Short-term: on Monday, set Tuesday's **staffing** levels (or next week's)
- **Design**: physician protocol – explore (triage) vs. exploit (discharge)

Technion IE&M research (OR/QS/IE) + **SEELab** technologies:

- **Operational recipes** (control, staffing, design) via Math-Nets (Q, F, D)
- **Virtual realities** (validation labs) via SEEnimations
- **Process Mining** (Discovery, Conformance) of ServNets
  (Started w/ **A. Gal, A. Senderovic, M. Weidlich**)

Feasible? Why be Optimistic?

e.g. Emergency-Department simulation, queueing model & analysis of a specific ED:

- Online: create **model(s)** of ED on normal Monday mornings
- Real-time: **control** patient-flow bottlenecks
- Short-term: on Monday, set Tuesday's **staffing** levels (or next week's)
- **Design**: physician protocol – explore (triage) vs. exploit (discharge)

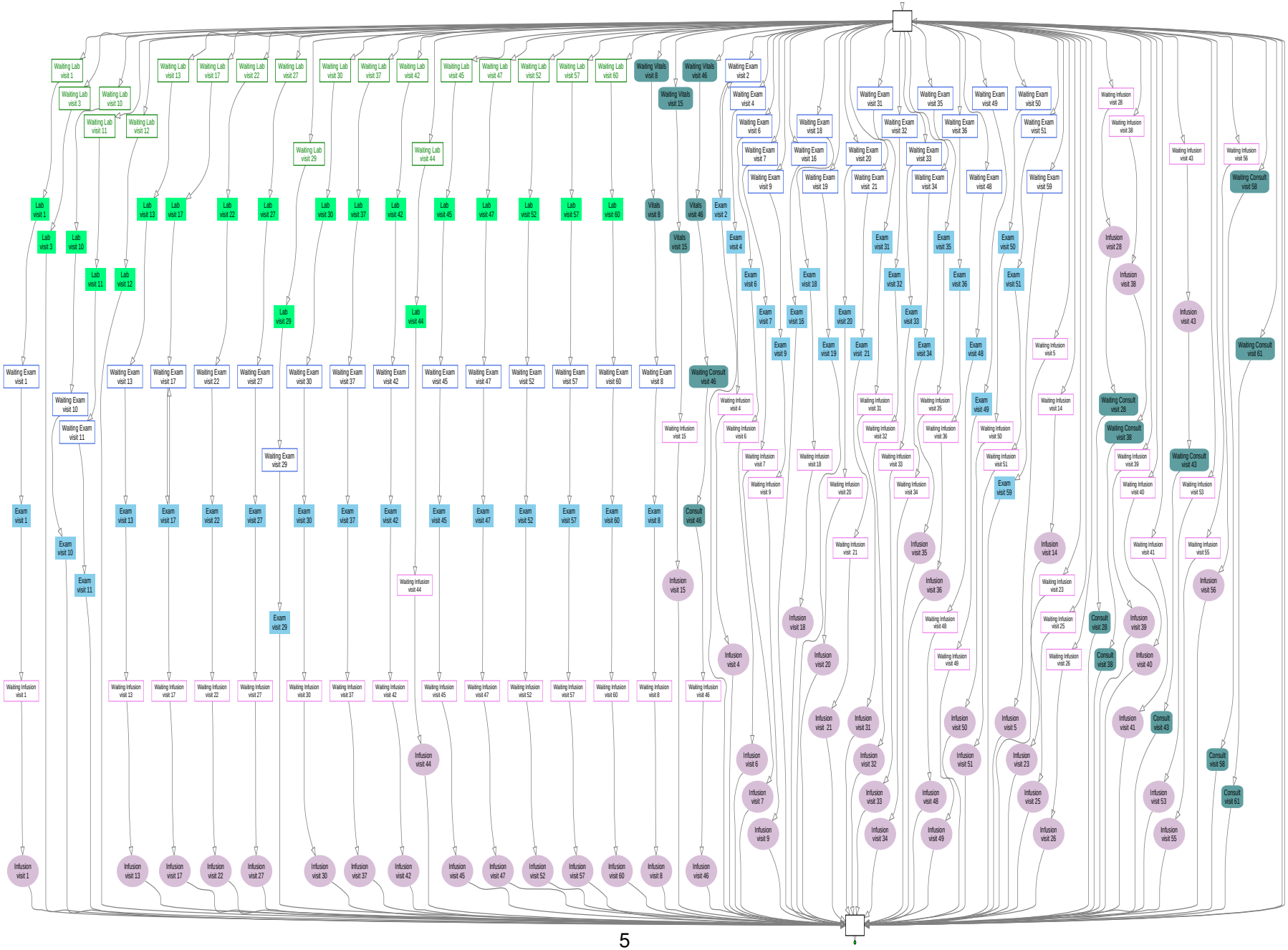
Technion IE&M research (OR/QS/IE) + **SEELab** technologies:

- **Operational recipes** (control, staffing, design) via Math-Nets (Q, F, D)
- **Virtual realities** (validation labs) via SEEnimations
- **Process Mining** (Discovery, Conformance) of ServNets
  (Started w/ **A. Gal, A. Senderovic, M. Weidlich**)

OR/SE and BPM/PM must join forces:

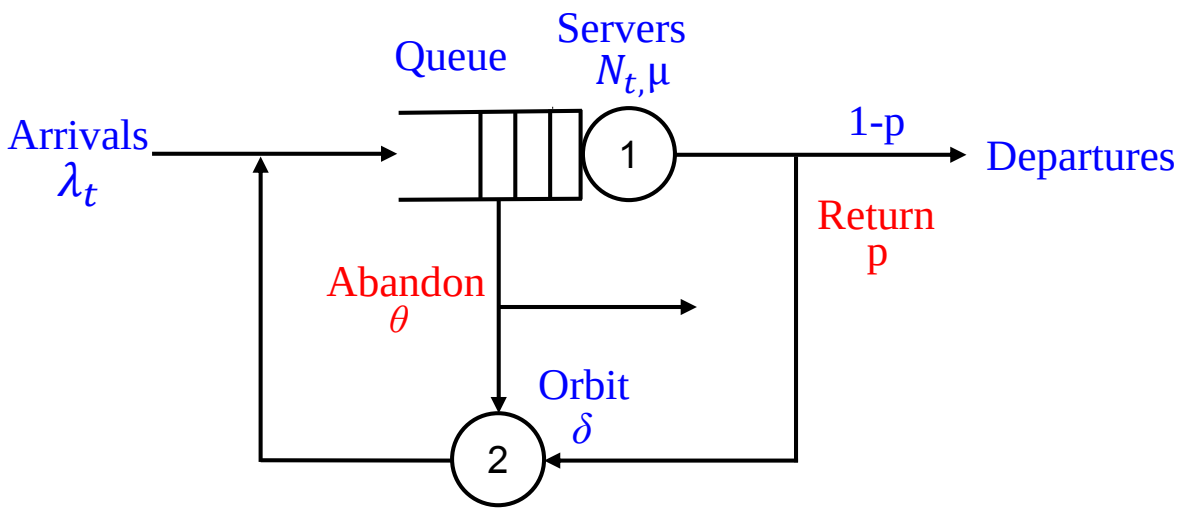
- OR/SE overcomes "**curse of dimensionality**": Simple models at the service of BPM/PM complex realities
- BPM/PM overcomes "**curse of parsimony**": OR models too simple to be credible

PM CarePath = 61 Visits within 6 months (DayHospital)



OR Model Selection: As Simple as Possible but Not Simpler

- Two-station ServNet, with customers who abandon and return
- Accommodates both **EDs and call centers** (**w/ Massey, Reiman, Stolyar**)



Contents

1. **Data-based research paradigm** (for me and interested colleagues): Operations Research, Industrial Engineering, Queueing Science

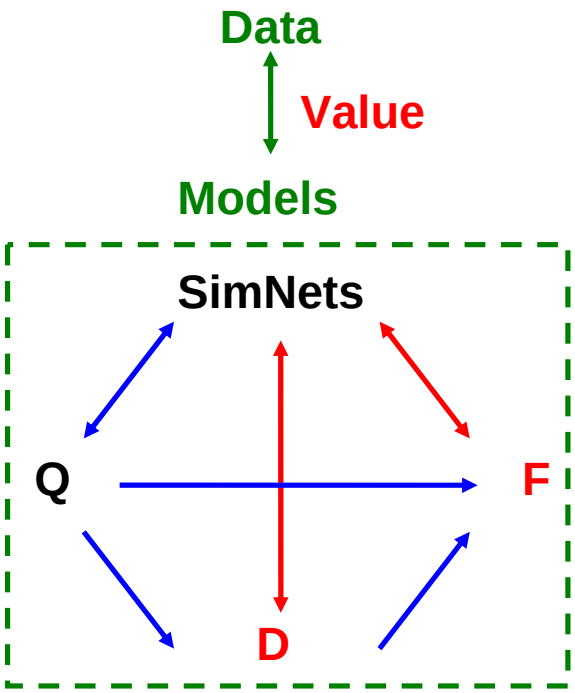
2. **Two simple (parsimonious) models**
   - Emergency Department: Time-Varying
   - Call Center: Stationary (“Fruit-Flies" of hospitals)

   (Reconciling time-variability and stationarity: stabilizing performance)

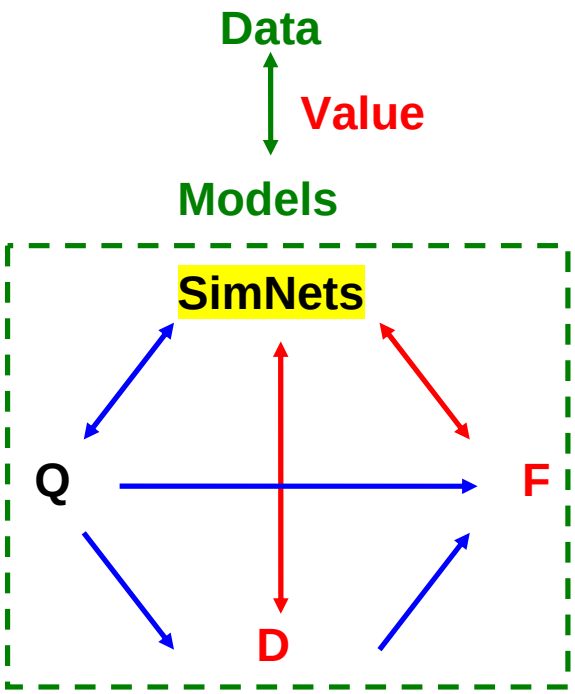
3. **Empirical adventures at the Technion IE&M SEELab**: Mining operational building blocks of ServNets
   - Primitives
   - Structure
   - Protocols

4. Glance at open problems, new directions, uncharted territories

Data–Based Framework: (Almost) All Models Born Equal



Data–Based (Asymptotic) Framework: Simulation Mining



Data-Based

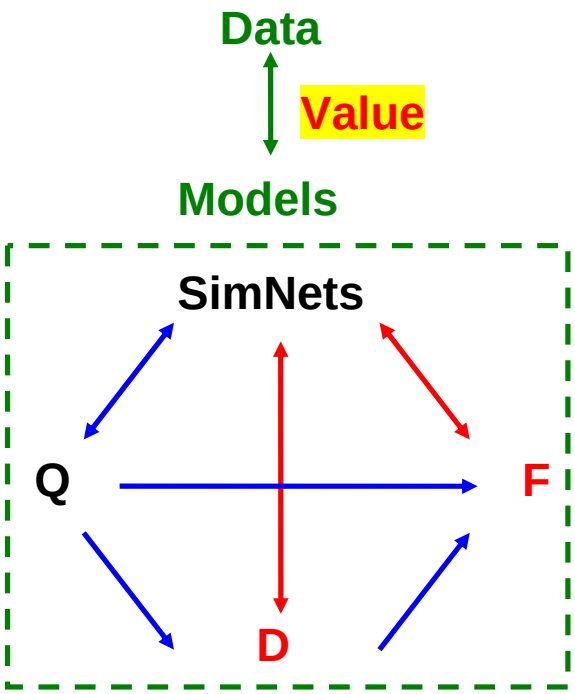
Real-Time

- Creation

- Application

- **“Implies” other ServNets**
- **Virtual Realities**
- **Validation: Parsimony**
- **Theory: Reproducibility**
- **Application: Trust**

Data–Based Asymptotic Framework: Added Value



Substitution Principle

Congestion Laws

Inference: Patience, Predictors

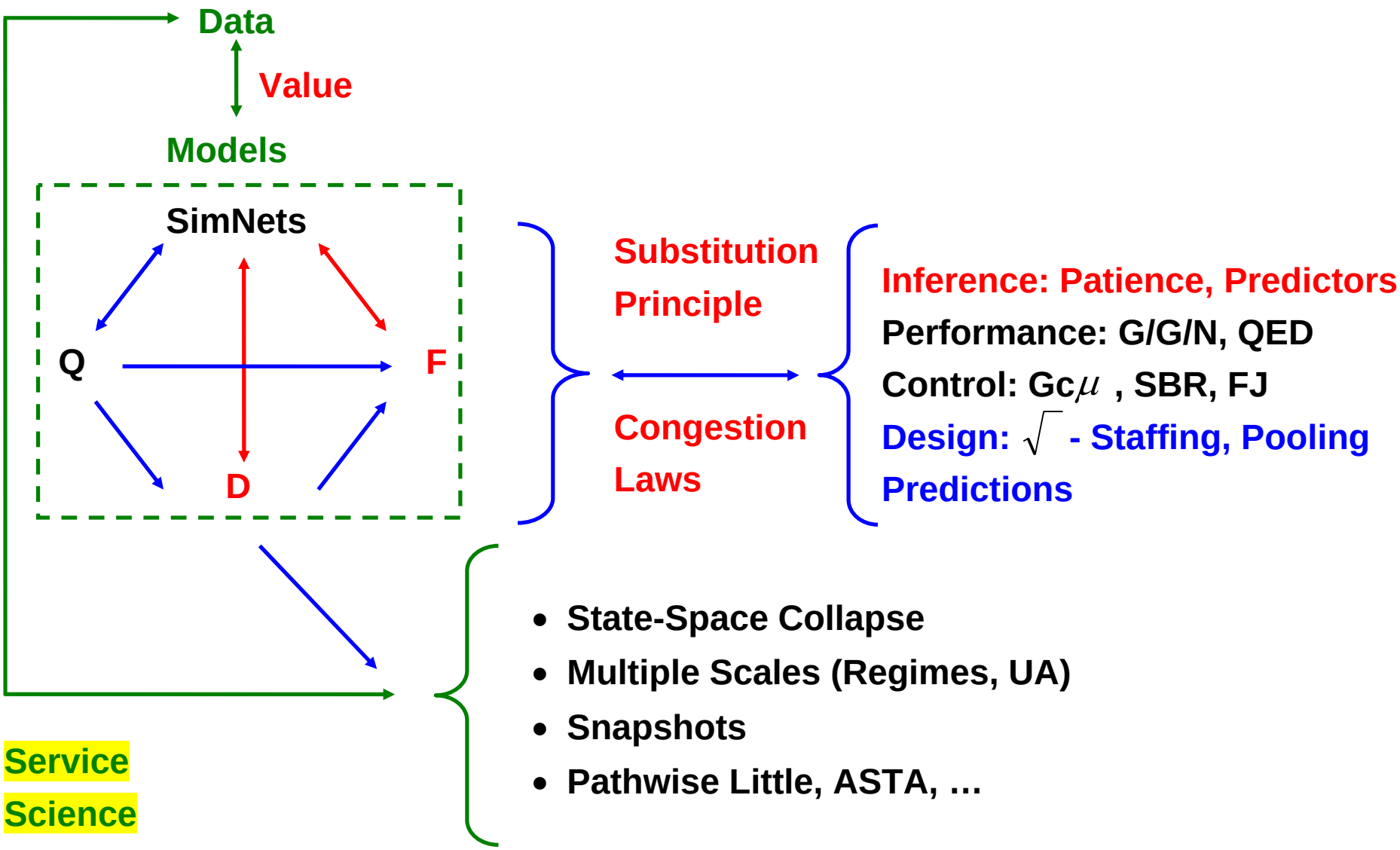
Performance: G/G/N, QED

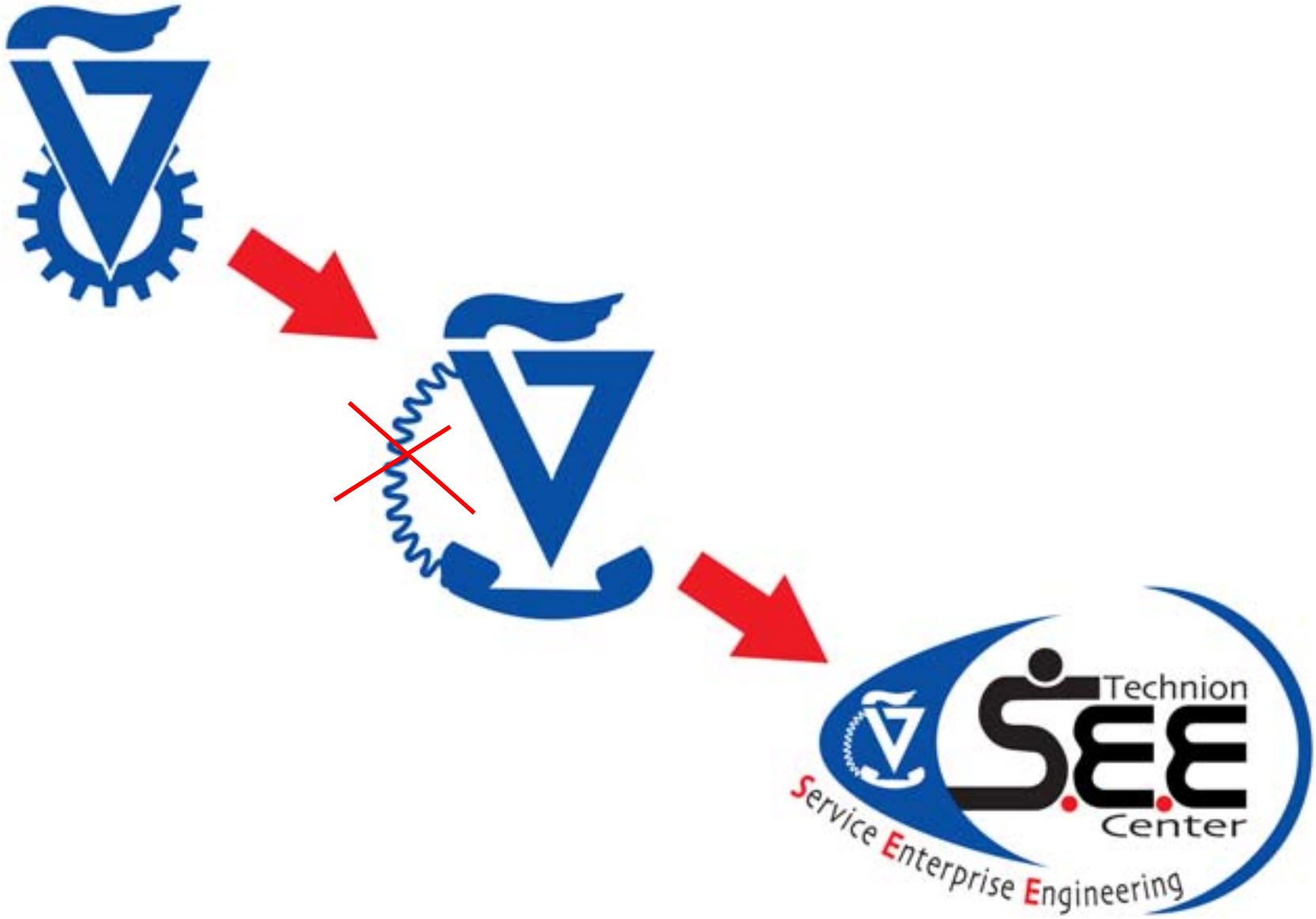
Control: Gcμ , SBR, FJ

Design: $\sqrt{}$ - Staffing, Pooling

Predictions

Ultimately: Automatic "Discovery, Conformance, Enhancement"




Technion
S.E.E
Center
Service Enterprise Engineering

Technion SEE = Service Enterprise Engineering

SEELab: Environment for graphical EDA in real-time

- Detailed **operational histories** (customers, servers), e.g.
  1. * Bank Anonymous: **1 year, 350K calls by 15 agents** - in 2000, which paved the way to:
  2. *U.S. Bank: **2.5 years, 220M calls, 40M by 1000 agents**
  3. Israeli Cellular: **2.5 years, 110M calls, 25M calls by 750 agents**
  4. Israeli Bank: **from January 2010, daily-deposit at a SEESafe**
  5. Service Engineering **internet site**: click-stream data (2 years)
  6. *Home (Rambam) Hospital: **4 years, 1000 beds**, inter-ward flow
  7. **Emergency Department** (ED) patient flow:
     - 5 EDs in Israel: 1-2 years, **late David Sinreich**, ED arrivals & LOS
     - ED in Seoul: 2 months, **K. Song-Hee & W. Cha**, pilot
     - ED in XY: 2 years, pilot
  8. Hospital **RTLS** (Real-Time Location System), pilot:
     - **250K events/day: 1000 patients, 350 staff (1500 tagged entities)**
     - Infrastructure: **900 readers (sensors), many floors**

*** Open & Free for research and teaching**

Data Cleaning: MCE with RFID Support

Data-base				Company report		comment
Asset id	order	**Entry** date	**Exit** date	Entry date	Exit date	
4	1	1:14:07 PM		1:14:00 PM		
6	1	12:02:02 PM	12:33:10 PM	12:02:00 PM	12:33:00 PM	
8	1	11:37:15 AM	12:40:17 PM	11:37:00 AM		**exit is missing**
10	1	12:23:32 PM	12:38:23 PM	12:23:00 PM		
12	1	12:12:47 PM	12:35:33 PM		12:35:00 PM	**entry is missing**
15	1	1:07:15 PM		1:07:00 PM		
16	1	11:18:19 AM	11:31:04 AM	11:18:00 AM	11:31:00 AM	
17	1	1:03:31 PM		1:03:00 PM		
18	1	1:07:54 PM		1:07:00 PM		
19	1	12:01:58 PM		12:01:00 PM		
20	1	11:37:21 AM	12:57:02 PM	11:37:00 AM	12:57:00 PM	
21	1	12:01:16 PM	12:37:16 PM	12:01:00 PM		
22	1	12:04:31 PM	12:20:40 PM			first customer is missing
22	2	12:27:37 PM		12:27:00 PM		
25	1	12:27:35 PM	1:07:28 PM	12:27:00 PM	1:07:00 PM	
27	1	12:06:53 PM		12:06:00 PM		
28	1	11:21:34 AM	11:41:06 AM	11:41:00 AM	11:53:00 AM	**exit time instead of entry time**
29	1	12:21:06 PM	12:54:29 PM	12:21:00 PM	12:54:00 PM	
31	1	11:40:54 AM	12:30:16 PM	11:40:00 AM	12:30:00 PM	
31	2	12:37:57 PM	12:54:51 PM	12:37:00 PM	12:54:00 PM	
32	1	11:27:11 AM	12:15:17 PM	11:27:00 AM	12:15:00 PM	
33	1	12:05:50 PM	12:13:12 PM	12:05:00 PM	12:15:00 PM	wrong exit time
35	1	11:31:48 AM	11:40:50 AM	11:31:00 AM	11:40:00 AM	
36	1	12:06:23 PM	12:29:30 PM	12:06:00 PM	12:29:00 PM	
37	1	11:31:50 AM	11:48:18 AM	11:31:00 AM	11:48:00 AM	
37	2	12:59:21 PM		12:59:00 PM		

- Imagine **“Cleaning" 60,000+ customers per day** (call centers) !
- “Psychology" of Data Trust and Transfer (e.g. 2 years till transfer)

Event-Logs in a Call Center (Bank Anonymous)

A Data Sample (Excel worksheet)

vru+line	call_id	customer_id	priority	type	date	vru_entry	vru_exit	vru_time	q_start	q_exit	q_time	outcome	ser_start	ser_exit	ser_time	server
AA0101	44749	27644400	2	PS	990901	11:45:33	11:45:39	6	11:45:39	11:46:58	79	AGENT	11:46:57	11:51:00	243	DORIT
AA0101	44750	12887816	1	PS	990905	14:49:00	14:49:06	6	14:49:06	14:53:00	234	AGENT	14:52:59	14:54:29	90	ROTH
AA0101	44967	58660291	2	PS	990905	14:58:42	14:58:48	6	14:58:48	15:02:31	223	AGENT	15:02:31	15:04:10	99	ROTH
AA0101	44968	0	0	NW	990905	15:10:17	15:10:26	9	15:10:26	15:13:19	173	HANG	00:00:00	00:00:00	0	NO_SERVER
AA0101	44969	63193346	2	PS	990905	15:22:07	15:22:13	6	15:22:13	15:23:21	68	AGENT	15:23:20	15:25:25	125	STEREN
AA0101	44970	0	0	NW	990905	15:31:33	15:31:47	14	00:00:00	00:00:00	0	AGENT	15:31:45	15:34:16	151	STEREN
AA0101	44971	41630443	2	PS	990905	15:37:29	15:37:34	5	15:37:34	15:38:20	46	AGENT	15:38:18	15:40:56	158	TOVA
AA0101	44972	64185333	2	PS	990905	15:44:32	15:44:37	5	15:44:37	15:47:57	200	AGENT	15:47:56	15:49:02	66	TOVA
AA0101	44973	3.06E+08	1	PS	990905	15:53:05	15:53:11	6	15:53:11	15:56:39	208	AGENT	15:56:38	15:56:47	9	MORIAH
AA0101	44974	74780917	2	NE	990905	15:59:34	15:59:40	6	15:59:40	16:02:33	173	AGENT	16:02:33	16:26:04	1411	ELI
AA0101	44975	55920755	2	PS	990905	16:07:46	16:07:51	5	16:07:51	16:08:01	10	HANG	00:00:00	00:00:00	0	NO_SERVER
AA0101	44976	0	0	NW	990905	16:11:38	16:11:48	10	16:11:48	16:11:50	2	HANG	00:00:00	00:00:00	0	NO_SERVER
AA0101	44977	33689787	2	PS	990905	16:14:27	16:14:33	6	16:14:33	16:14:54	21	HANG	00:00:00	00:00:00	0	NO_SERVER
AA0101	44978	23817067	2	PS	990905	16:19:11	16:19:17	6	16:19:17	16:19:39	22	AGENT	16:19:38	16:21:57	139	TOVA
AA0101	44764	0	0	PS	990901	15:03:26	15:03:36	10	00:00:00	00:00:00	0	AGENT	15:03:35	15:06:36	181	ZOHARI
AA0101	44765	25219700	2	PS	990901	15:14:46	15:14:51	5	15:14:51	15:15:10	19	AGENT	15:15:09	15:17:00	111	SHARON
AA0101	44766	0	0	PS	990901	15:25:48	15:26:00	12	00:00:00	00:00:00	0	AGENT	15:25:59	15:28:15	136	ANAT
AA0101	44767	58859752	2	PS	990901	15:34:57	15:35:03	6	15:35:03	15:35:14	11	AGENT	15:35:13	15:35:15	2	MORIAH
AA0101	44768	0	0	PS	990901	15:46:30	15:46:39	9	00:00:00	00:00:00	0	AGENT	15:46:38	15:51:51	313	ANAT
AA0101	44769	78191137	2	PS	990901	15:56:03	15:56:09	6	15:56:09	15:56:28	19	AGENT	15:56:28	15:59:02	154	MORIAH
AA0101	44770	0	0	PS	990901	16:14:31	16:14:46	15	00:00:00	00:00:00	0	AGENT	16:14:44	16:16:02	78	BENSION
AA0101	44771	0	0	PS	990901	16:38:59	16:39:12	13	00:00:00	00:00:00	0	AGENT	16:39:11	16:43:35	264	VICKY
AA0101	44772	0	0	PS	990901	16:51:40	16:51:50	10	00:00:00	00:00:00	0	AGENT	16:51:49	16:53:52	123	ANAT
AA0101	44773	0	0	PS	990901	17:02:19	17:02:28	9	00:00:00	00:00:00	0	AGENT	17:02:28	17:07:42	314	VICKY
AA0101	44774	32387482	1	PS	990901	17:18:18	17:18:24	6	17:18:24	17:19:01	37	AGENT	17:19:00	17:19:35	35	VICKY
AA0101	44775	0	0	PS	990901	17:38:53	17:39:05	12	00:00:00	00:00:00	0	AGENT	17:39:04	17:40:43	99	TOVA

- Unsynchronized transition times, consistently

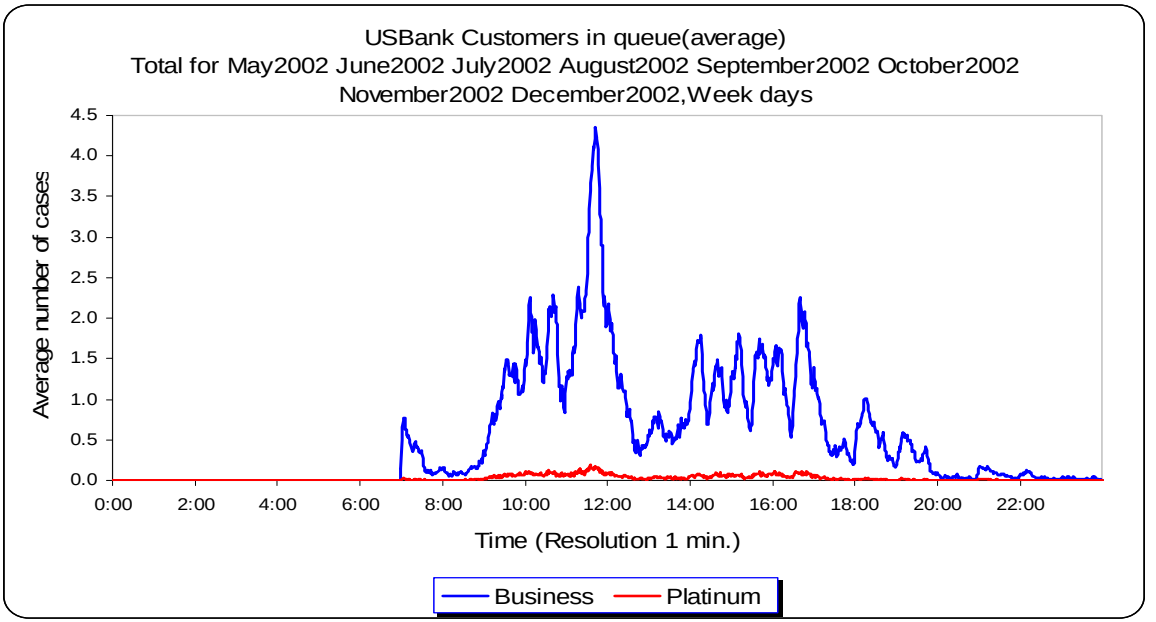
Technion - Israel Institute of Technology
The William Davidson Faculty of Industrial Engineering and Management

Center for Service Enterprise Engineering (SEE)
http://ie.technion.ac.il/Labs/Serveng/

SEEStat 3.0 Tutorial

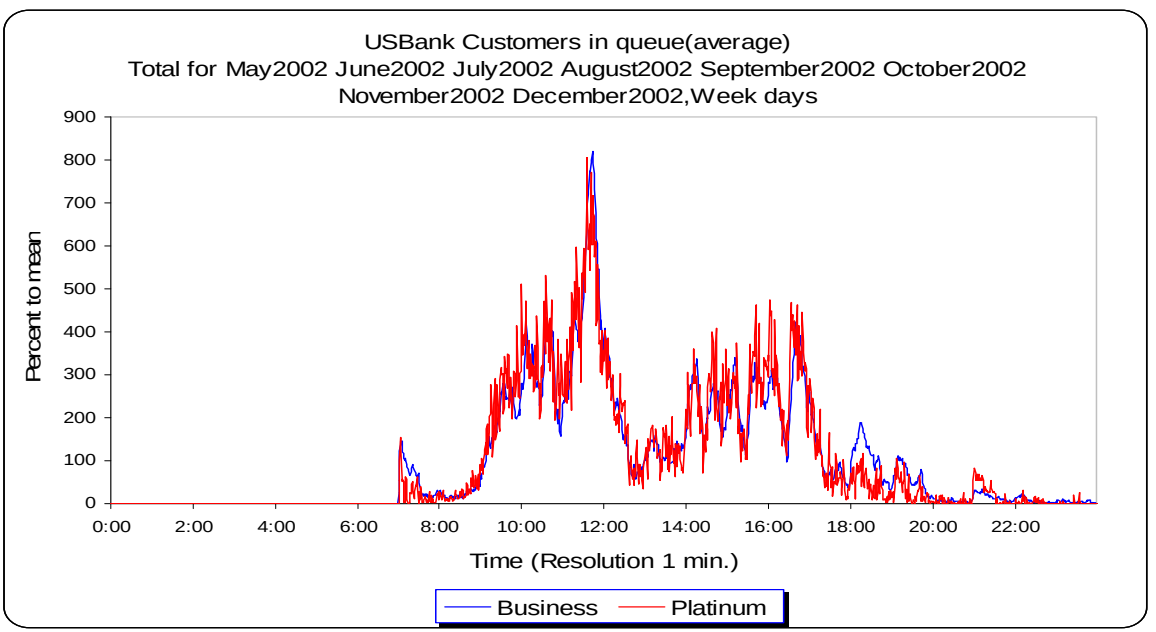
NYU, Stern IOMS, September 2011

Note: **This tutorial is customized to NYU Stern.**
To become a regular user of SEEStat, please go to
http://seeserver.iem.technion.ac.il/see-terminal/
click on “Register” (left menu), and follow the registration procedure.



Platinum is a small-scale service. You will now normalize the chart in order to identify patters.

Click **"Output"** on the main menu and then **"Modify Tables and Charts"**.
Open the **"Options"** tab and select **Percent to mean**. Click **"OK"**.



Note the essentially overlapping patterns of the queue lengths of the two customer types. (This phenomenon is predicted by asymptotic analysis of queues in heavy traffic, where it is referred to as State-Space-Collapse.)

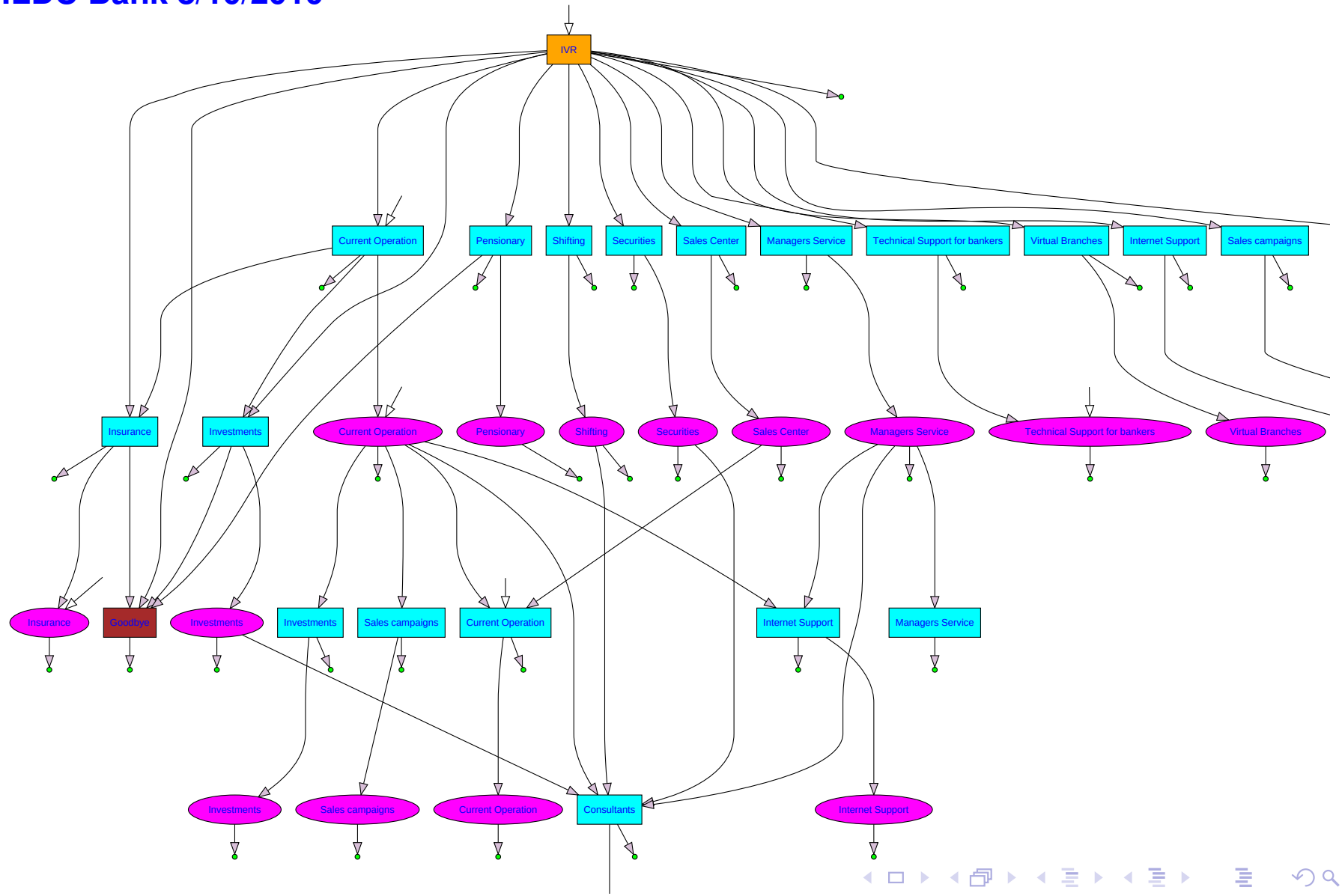
Call-Center Environment: Service Network

= "Fruit-flies of Hospitals": fast, low-stake, no IRB, . . .

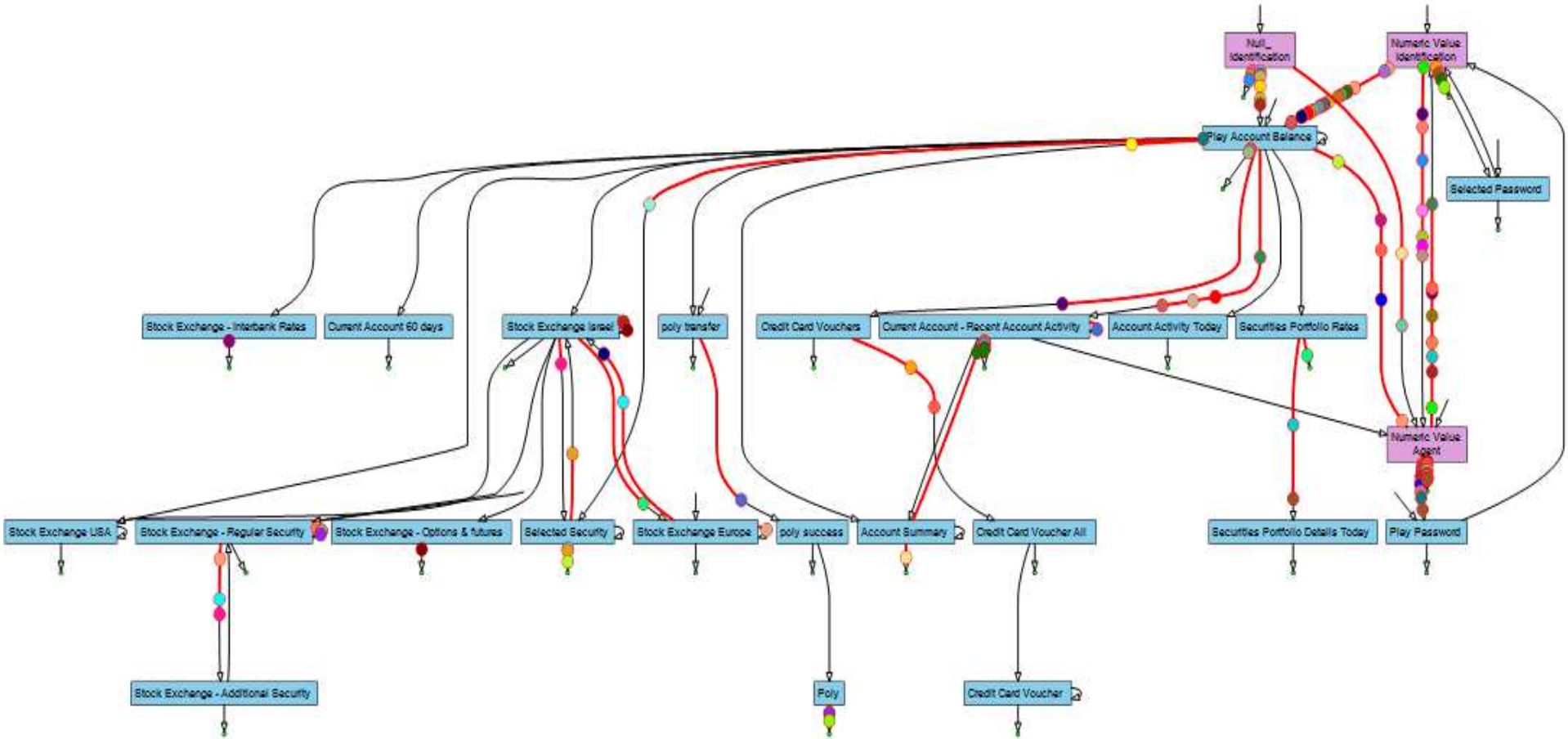
yet highly relevant

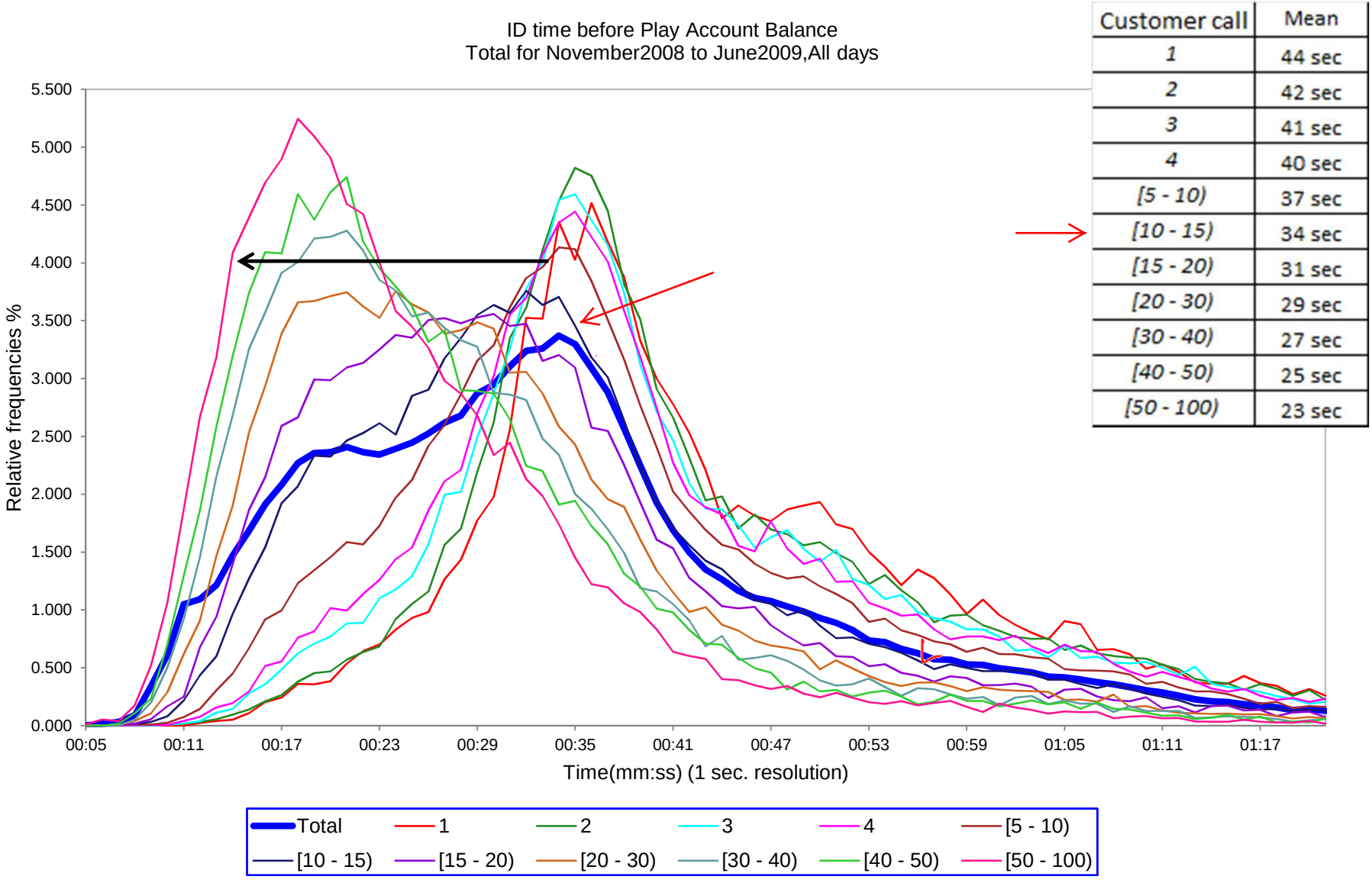
Customer Flow in Call Centers

ILDU Bank 8/10/2010



IVR flow

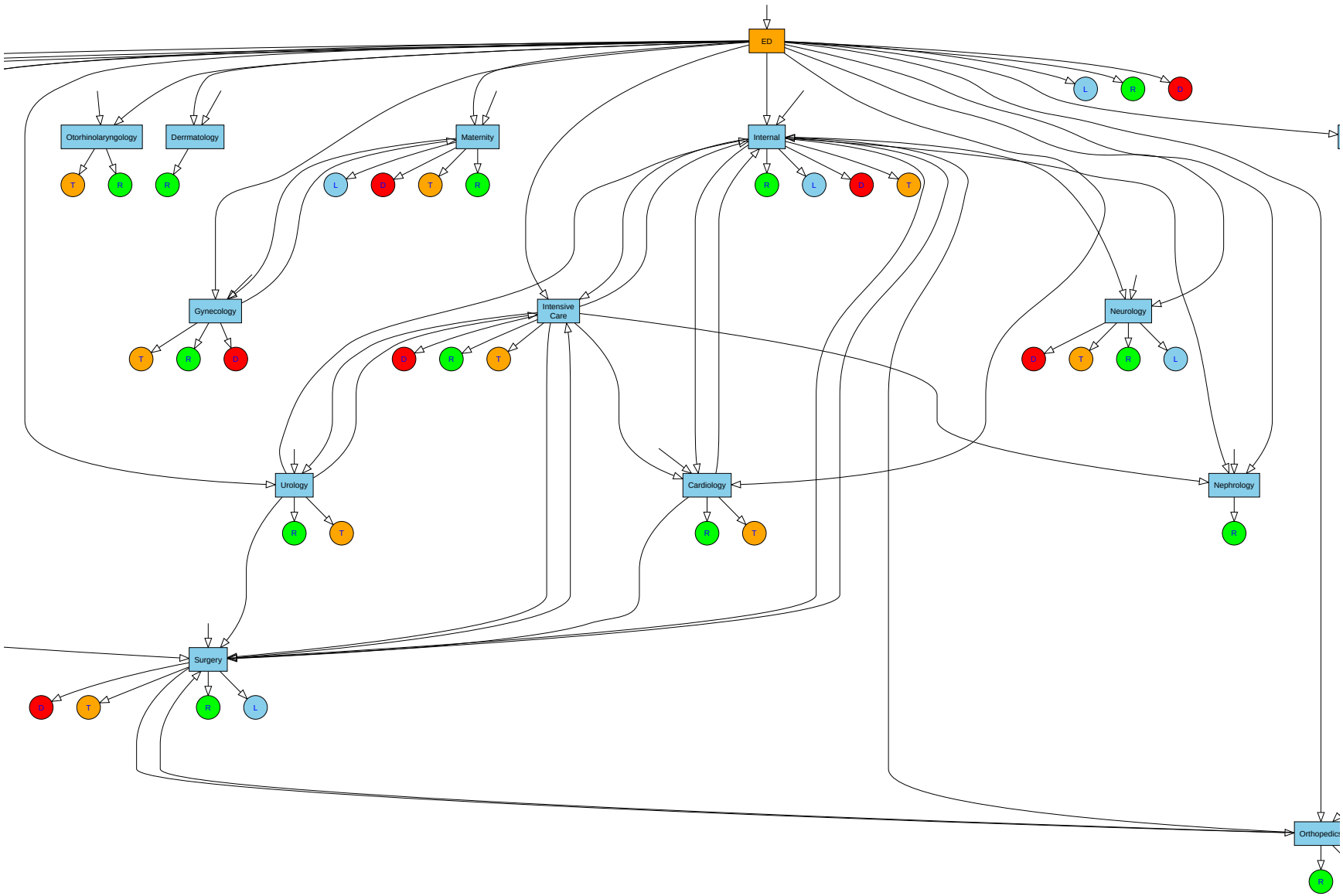

Null_ Identification
Numeric Value Identification
Play Account Balance
Selected Password
Stock Exchange - Interbank Rates
Current Account 60 days
Stock Exchange Israel
poly transfer
Credit Card Vouchers
Current Account - Recent Account Activity
Account Activity Today
Securities Portfolio Rates
Numeric Value Agent
Stock Exchange USA
Stock Exchange - Regular Security
Stock Exchange - Options & futures
Selected Security
Stock Exchange Europe
poly success
Account Summary
Credit Card Voucher All
Securities Portfolio Details Today
Play Password
Stock Exchange - Additional Security
Poly
Credit Card Voucher



Customer call	Mean
1	44 sec
2	42 sec
3	41 sec
4	40 sec
[5 - 10)	37 sec
[10 - 15)	34 sec
[15 - 20)	31 sec
[20 - 30)	29 sec
[30 - 40)	27 sec
[40 - 50)	25 sec
[50 - 100)	23 sec

Hospital Network: Inter-Ward Patient Flow

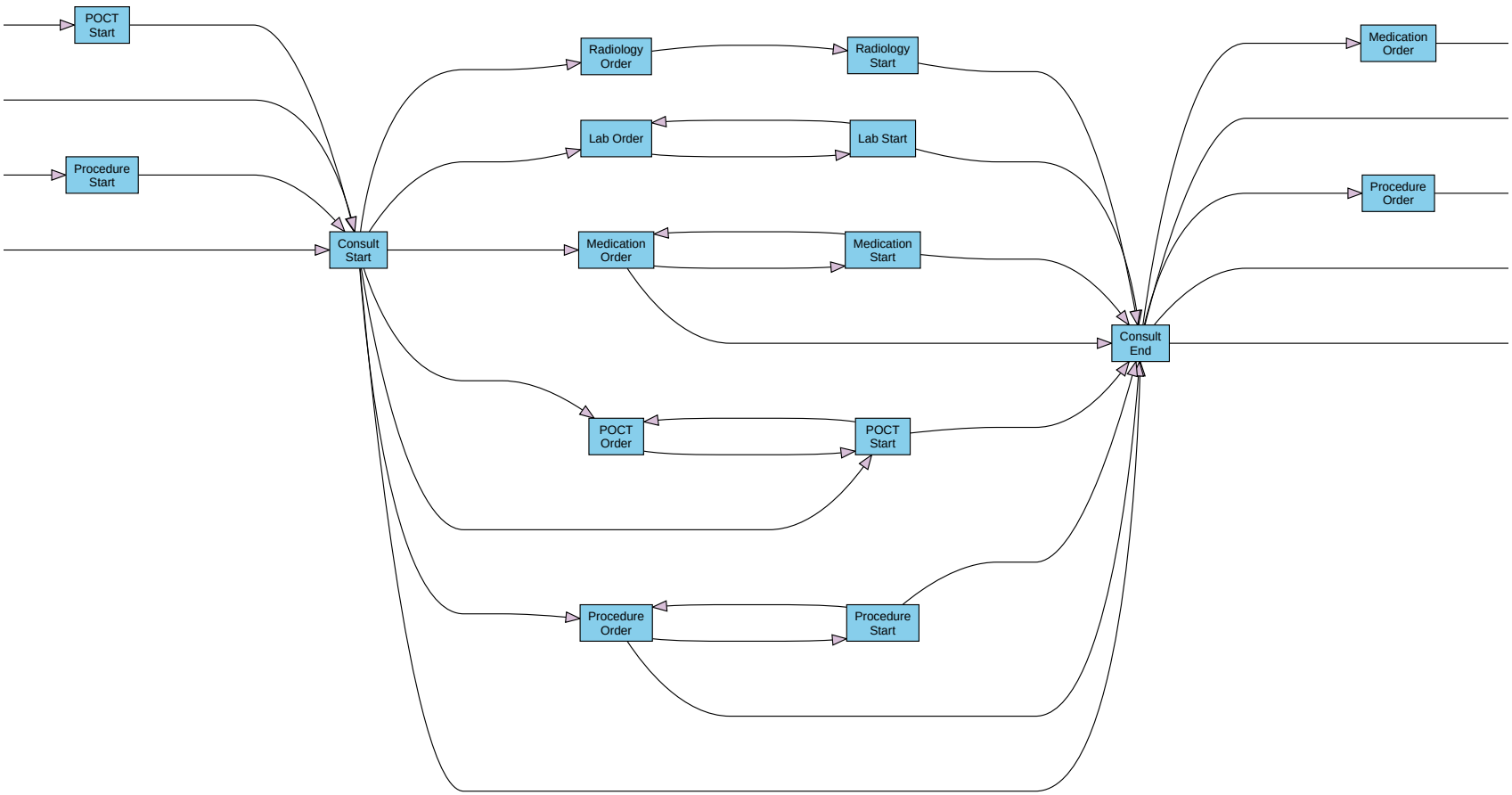
Patients movement (HomeHospital)
January 2005

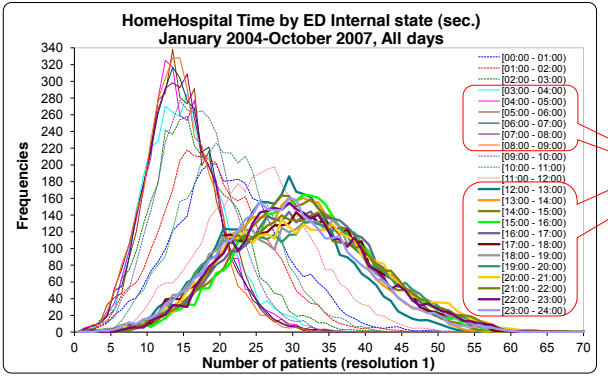


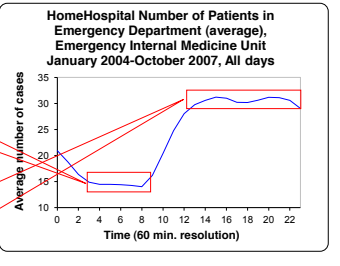
Simple Models at the Service of Complex Realities: FNets

1. ED in Normal days (Time-Varying Periodic): Personnel Staffing (offline)
2. ED in **Mass Casualty Event** (**Transient**): Forecasting, Staffing (online)

Emergency Department in XYHospital, October 2012

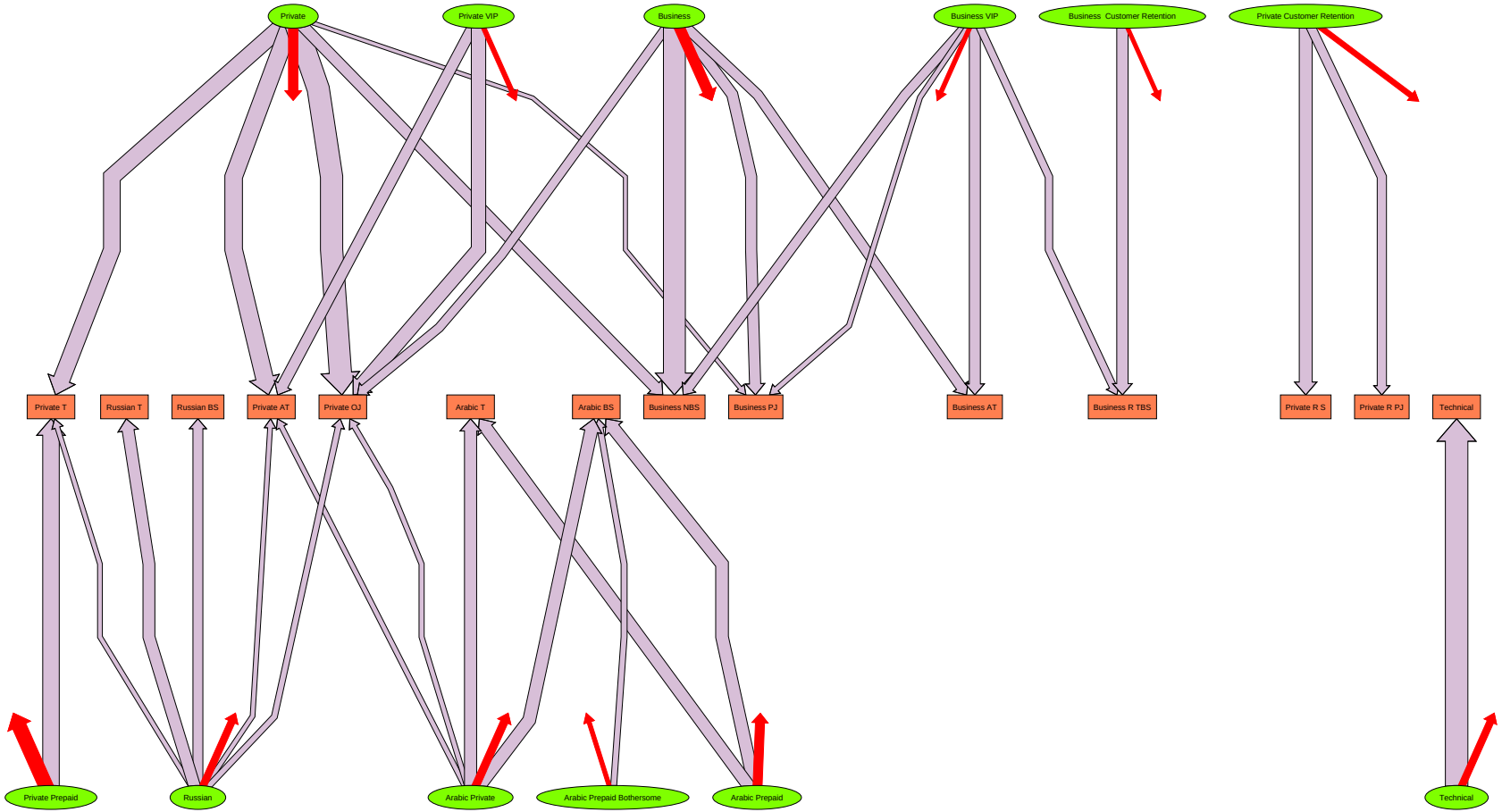






Internal ED Occupancy histogram (left) and Average Census (right), by hour of the day

ILTelecom 9/3/2008



Empirical Adventures at the SEELab

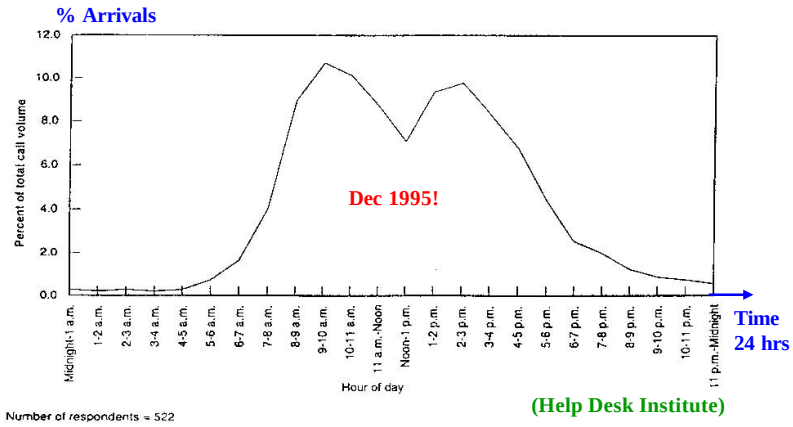
SEELab History suggests possible guidelines for ServNet Mining:

1. **Primitives**: arrivals, services, (im)patience
2. **Structure**: static process-maps
3. **Protocols**: Load Balancing, Dynamic Priority, Information

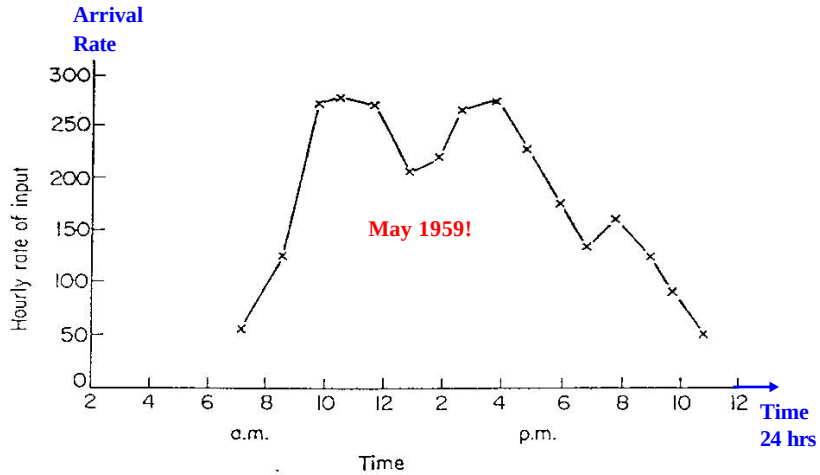
Primitives: Arrival (Rates) to Service

Why 2 Daily Peaks?

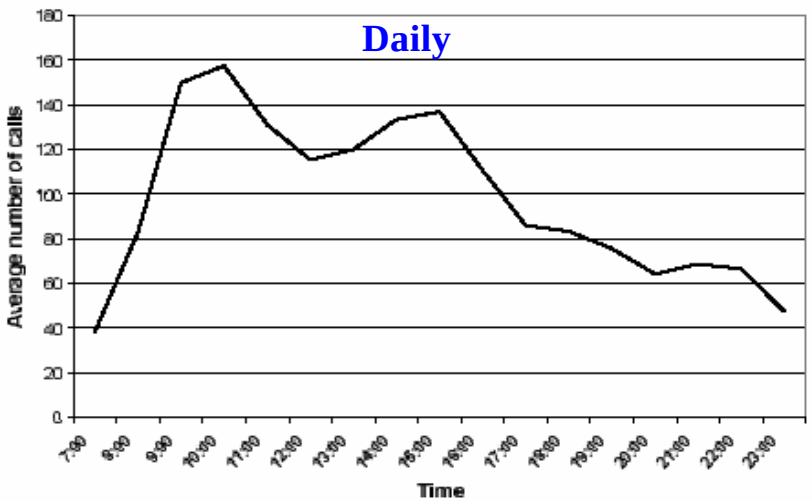
CC: Dec. **1995**, (USA, 700 Helpdesks)



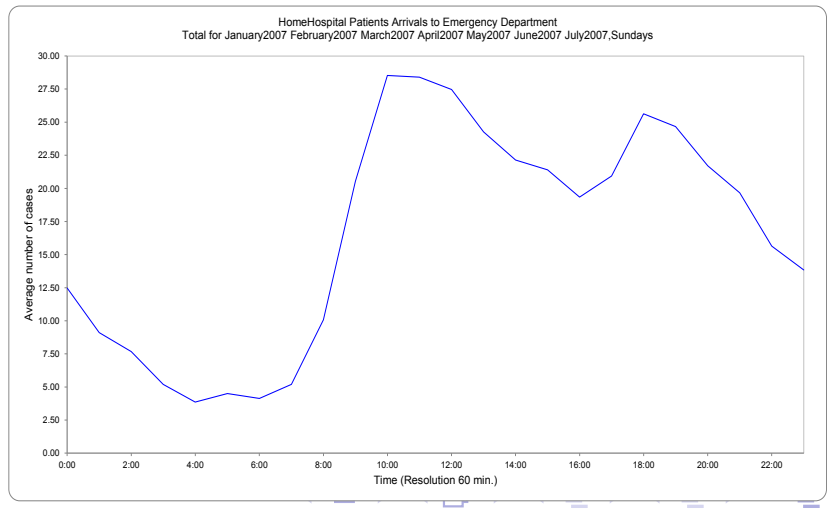
CC: May **1959** (England)



CC: Nov. 1999 (Israel)

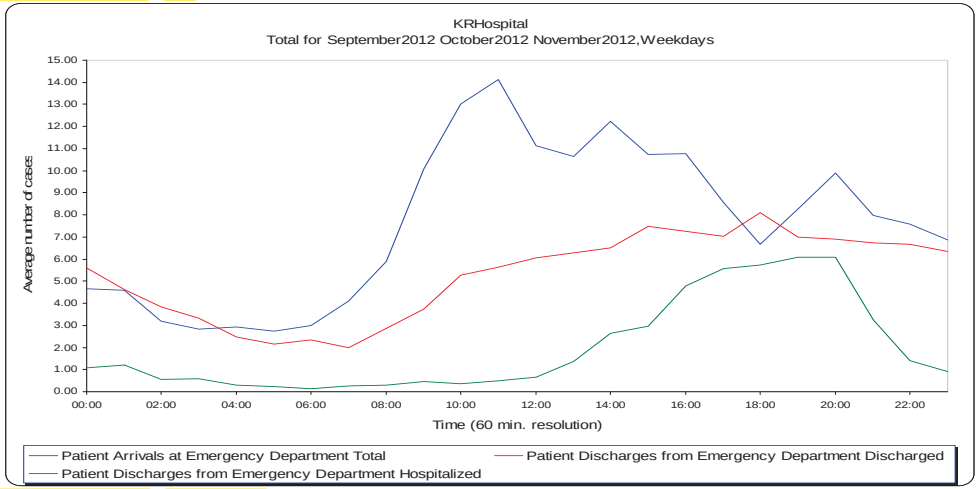


ED: Jan.–July 2007 (Israel)

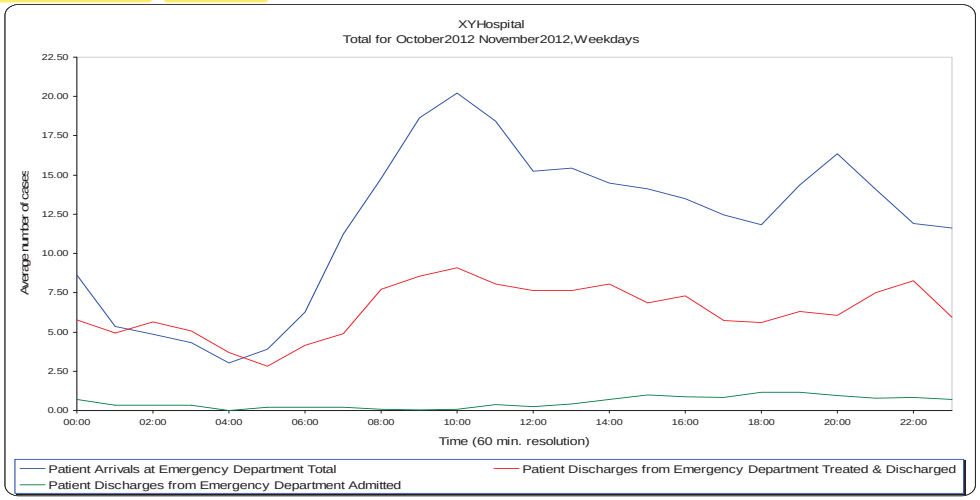


Arrival (Discharge) Rates in Korea and Singapore

KRHospital, all ED patients



XYHospital, walking patients

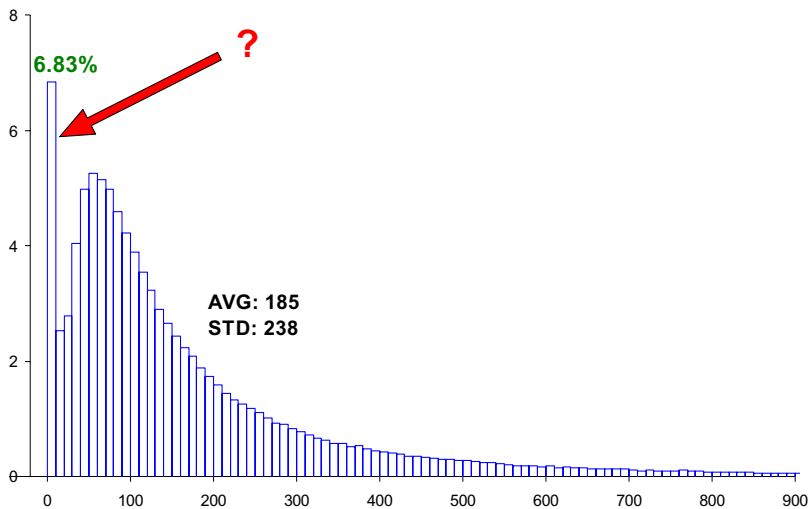


Primitives: Services (Durations)

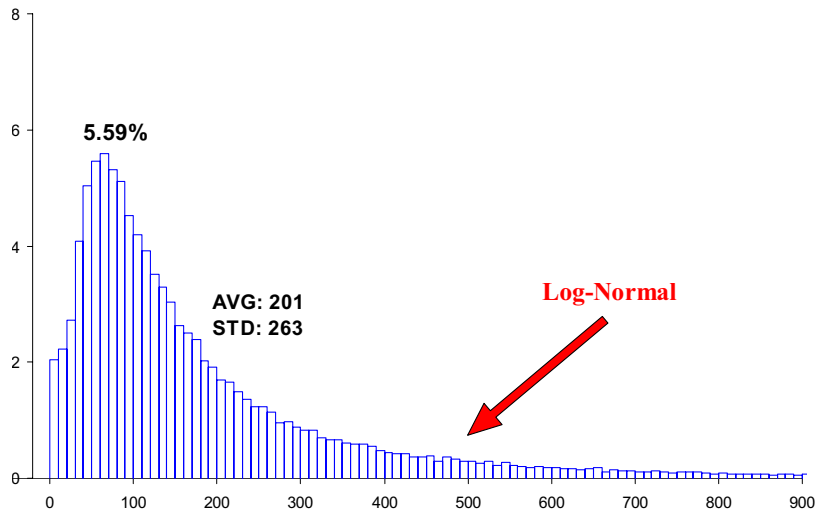
Histogram of Service-Duration in an Israeli Call Center, 1999

Why short services? Why LogNormal?

January-October



November-December

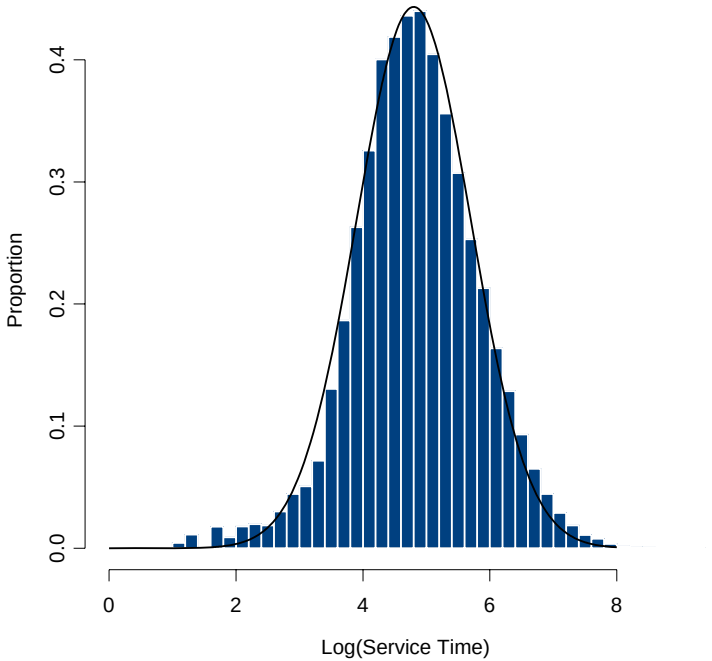


- January-October: **6.8% Short-Services** ($\leq$ 10 seconds) ?
- November-December: **LogNormal** durations (common) ?

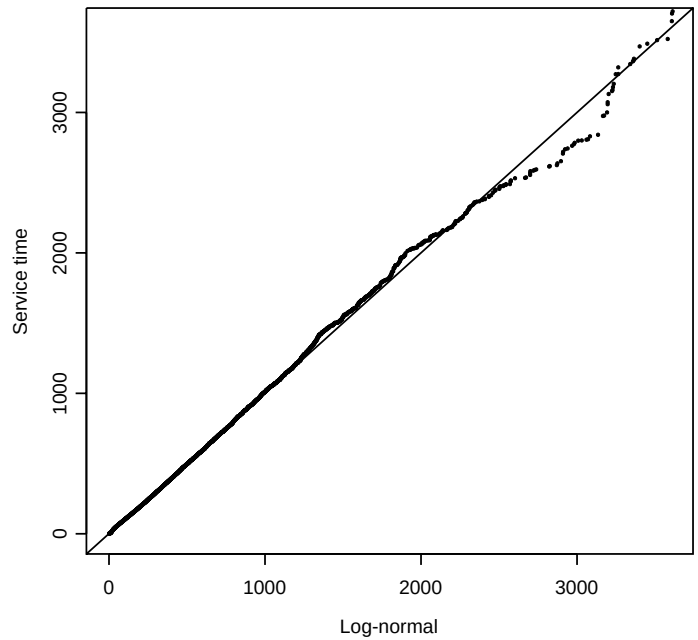
Durations: Phone Calls (2 Surprises)

Israeli Call Center, Nov–Dec, 1999

Log(Service Times)



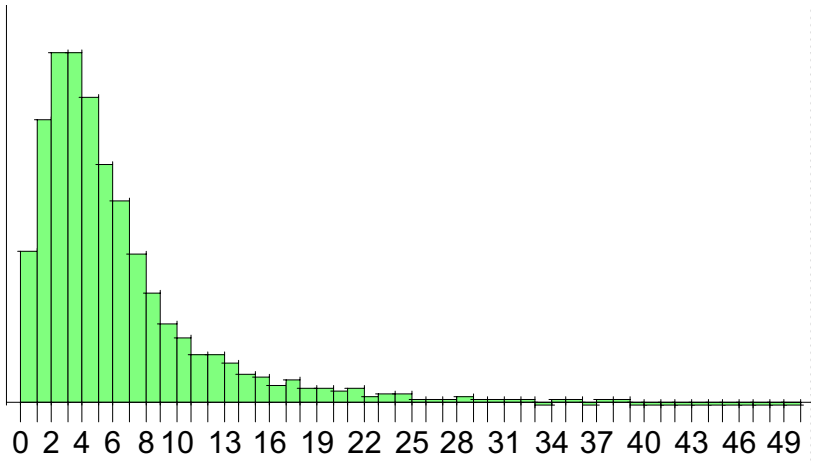
LogNormal QQPlot



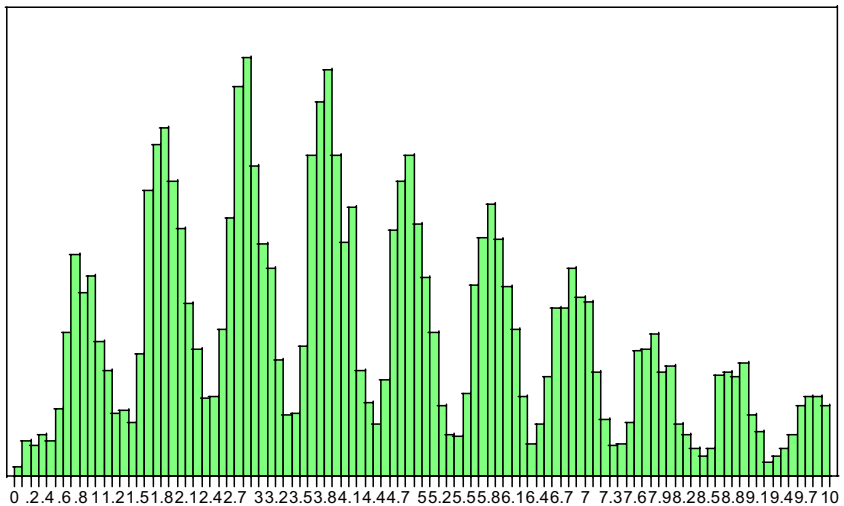
- **Practically Important**: (mean, std)(log) characterization
- **Theoretically Intriguing**: Why LogNormal ? Naturally multiplicative but, in fact, also **Infinitely-Divisible** (Generalized Gamma-Convolutions)

Protocols: LOS in Hospitals - Beyond LogNormal

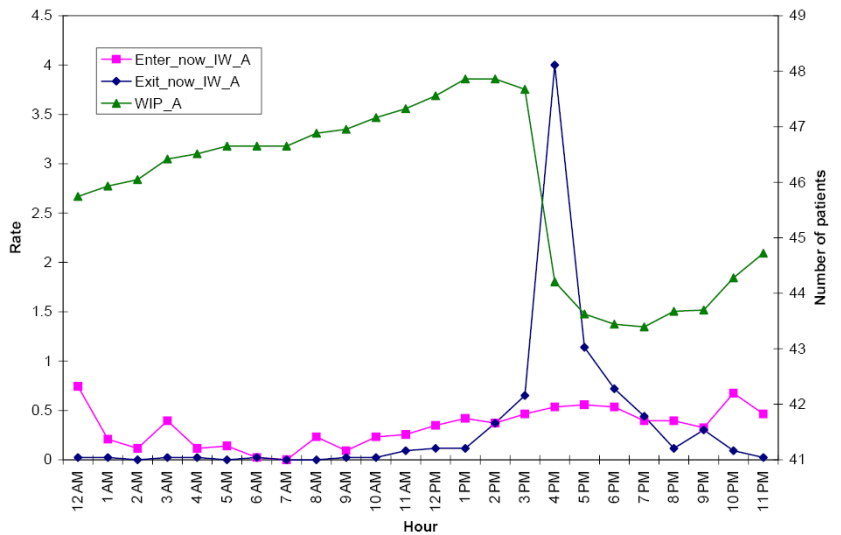
Israeli Hospital, in Days: LN



In Hours: 2 Time Scales, Mixture



Explanation: Patients released around **3pm** (2-3 in Singapore, 2-4 in UNC Hospital)



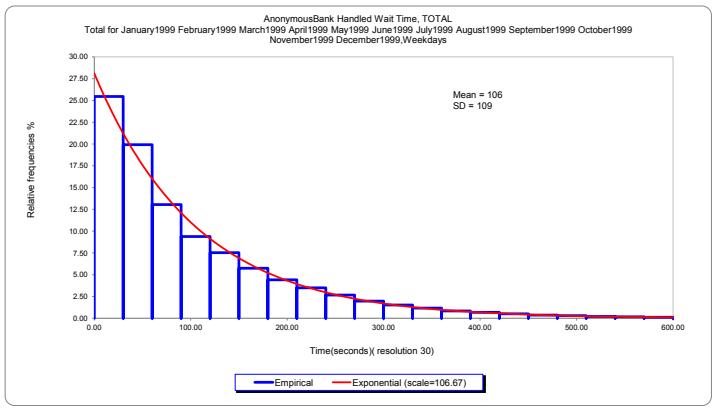
Why Bother ?

- Hourly Scale: Staffing,. . .
- Daily: Flow / Bed Control,. . .

Protocols: Waiting Time in a Call Center

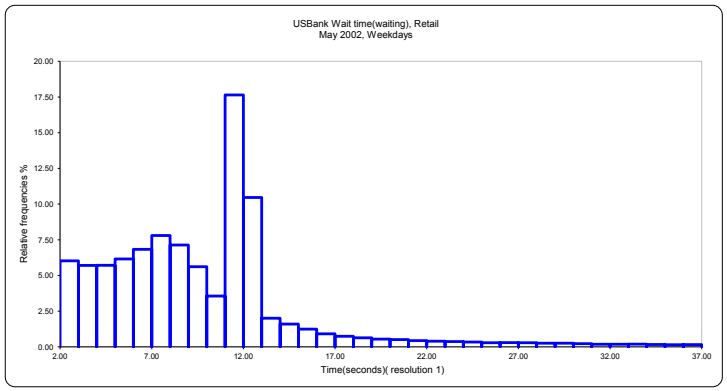
Exponential in Heavy-Traffic (min.)

Small Israeli Bank



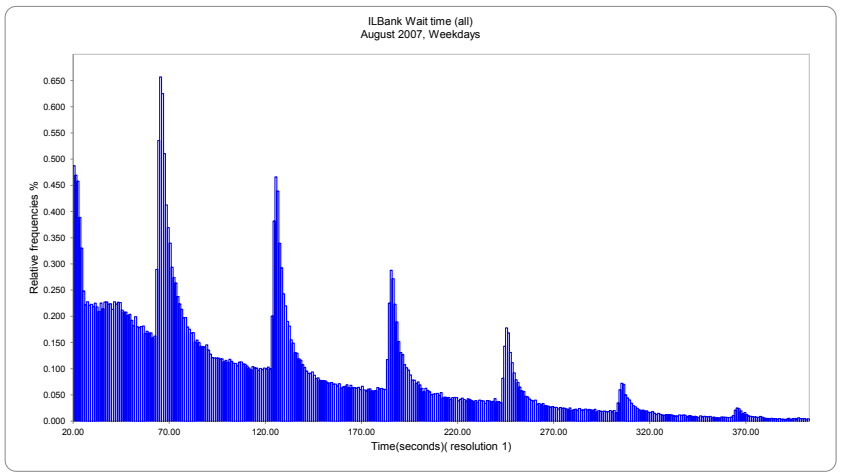
Routing via Thresholds (sec.)

Large U.S. Bank



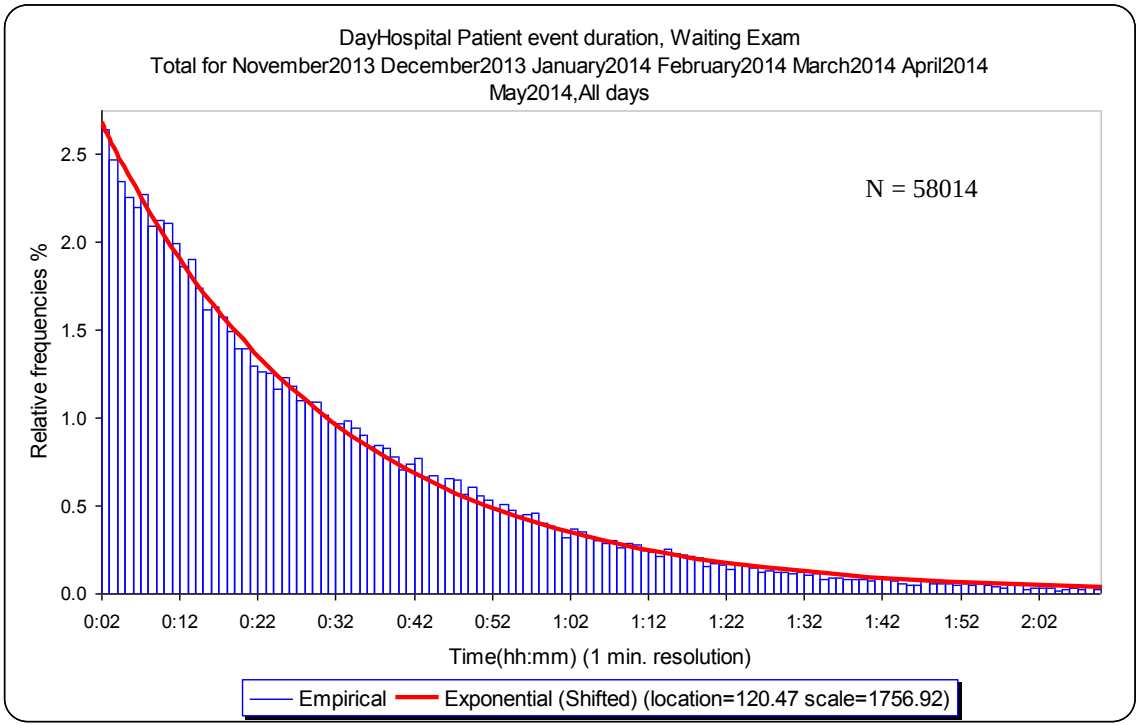
Scheduling Priorities (sec.) [compare Hospital LOS (hours)]

Medium Israeli Bank



Exam

Fitting distribution for waiting-time to exam



DayHospital Patient event duration, Waiting Exam, Total for November2013 December2013 January2014 February2014 March2014 April2014 May2014,All days	
	Patient event duration
N	58014
N(average per day)	479.45
Mean	31 min 16 sec
Standard Deviation	30 min 1 sec
Median	22 min 39 sec
Minimum	2 min
Maximum	7 hours 21 min#

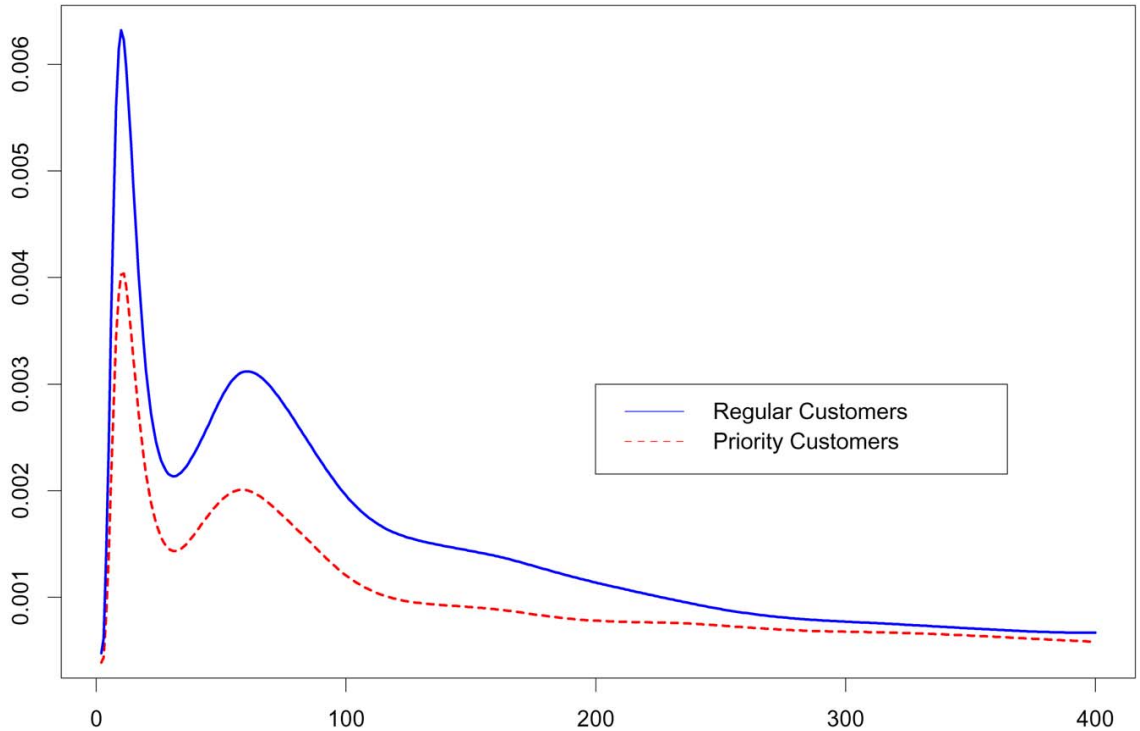
Parameters for Exponential (Shifted) Distribution	
Parameter	*Estimate*
location	120.47
scale	1756.92
mean	1877.4
std	1756.9

Goodness-of-Fit Tests for Exponential (Shifted) Distribution			
Test	*Statistic*	*DF*	*p Value*
Residuals Std	0.0046		
Kolmogorov-Smirnov	0.0083		0.001
Cramer-von Mises	1.2440		0.001
Anderson-Darling	9.8955		<.0001
Chi-Square	969.5762	20	<.0001

Protocols: (Im)Patience while Waiting (Psychology)

Palm: (1943–53): **Irritation** $\propto$ **Hazard Rate**

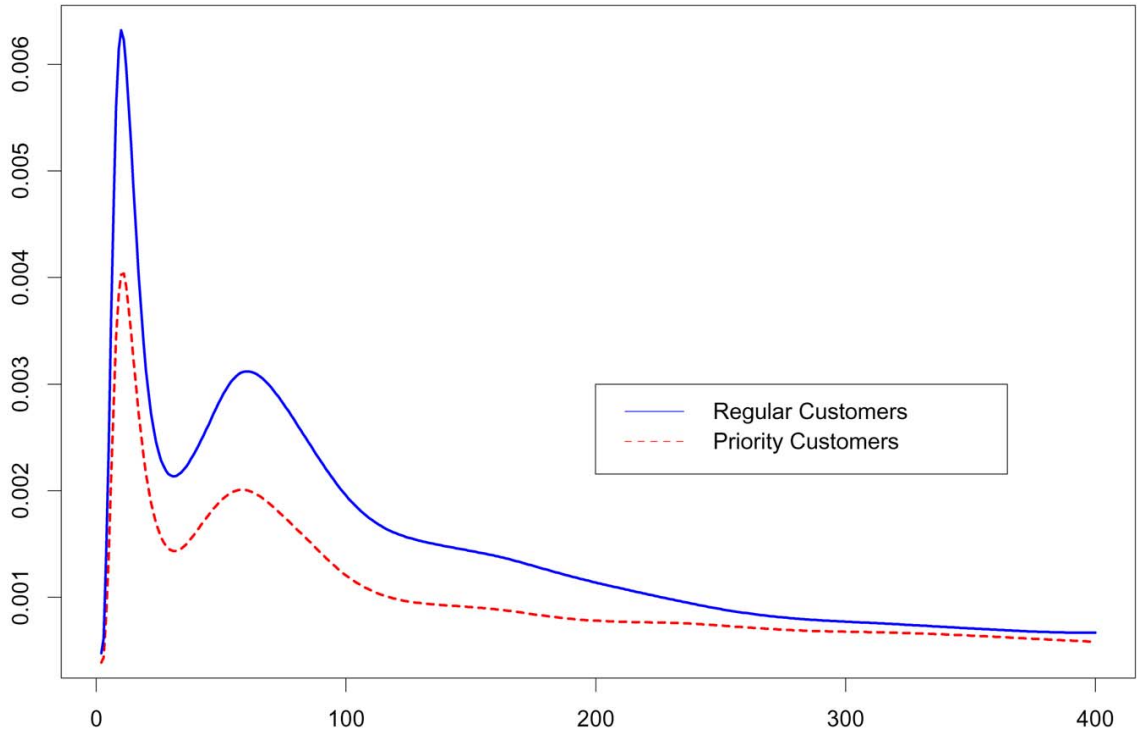
Regular over **VIP** Customers: VIP more patient here (Israeli Bank)



Protocols: (Im)Patience while Waiting (Psychology)

Palm: (1943–53): **Irritation** $\propto$ **Hazard Rate**

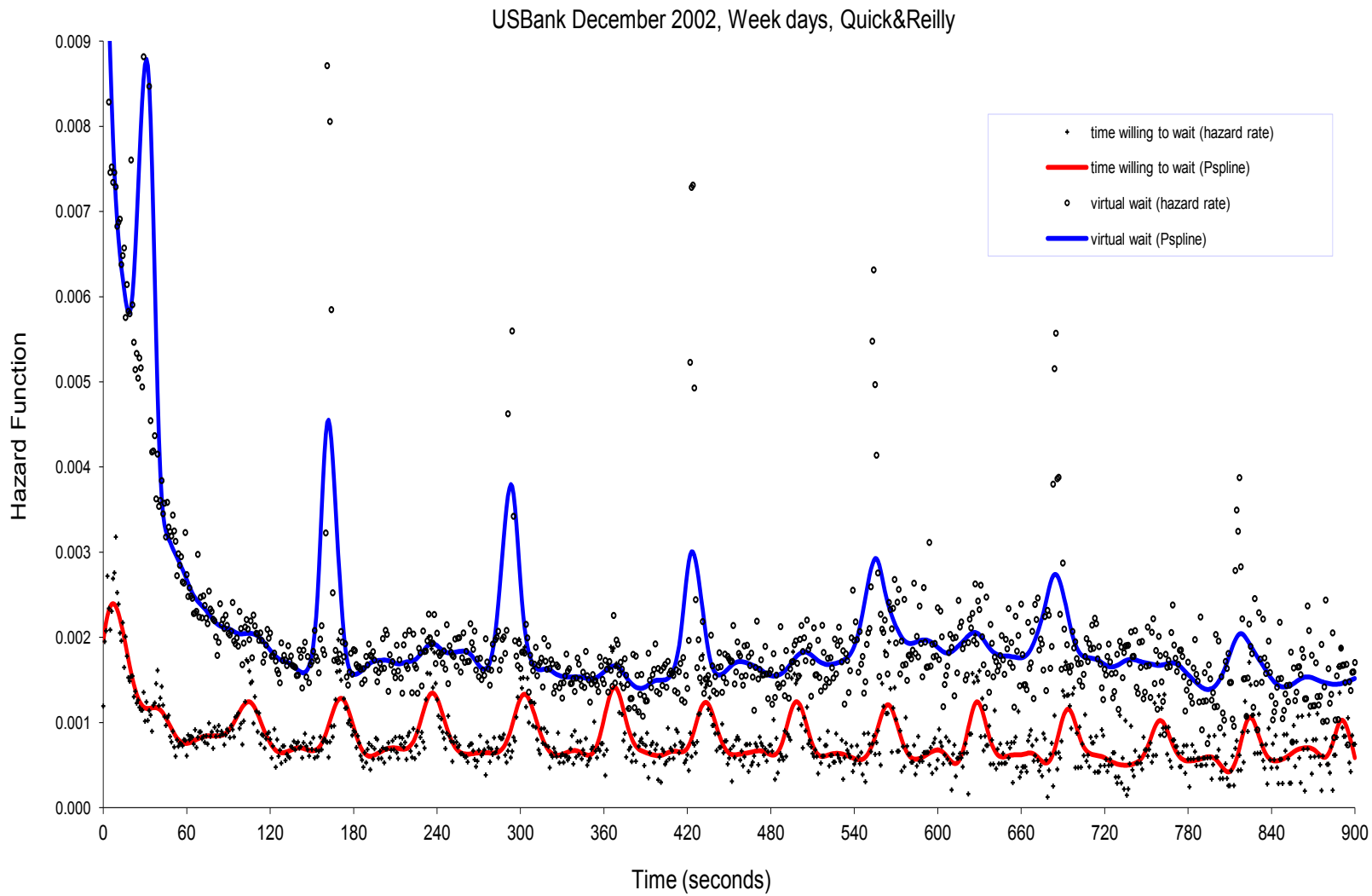
Regular over **VIP** Customers: VIP more patient here (Israeli Bank)



- Why **Peaks** of **abandonment**? **Announcement** epochs
  - Control abandonment w/ info: encourage, discourage
  - Technical Challenges, **w/ J. Huang, J. Zhang, H. Zhang**
- Statistical challenges: **Un-Censoring**, Smoothing

Protocols + Psychology

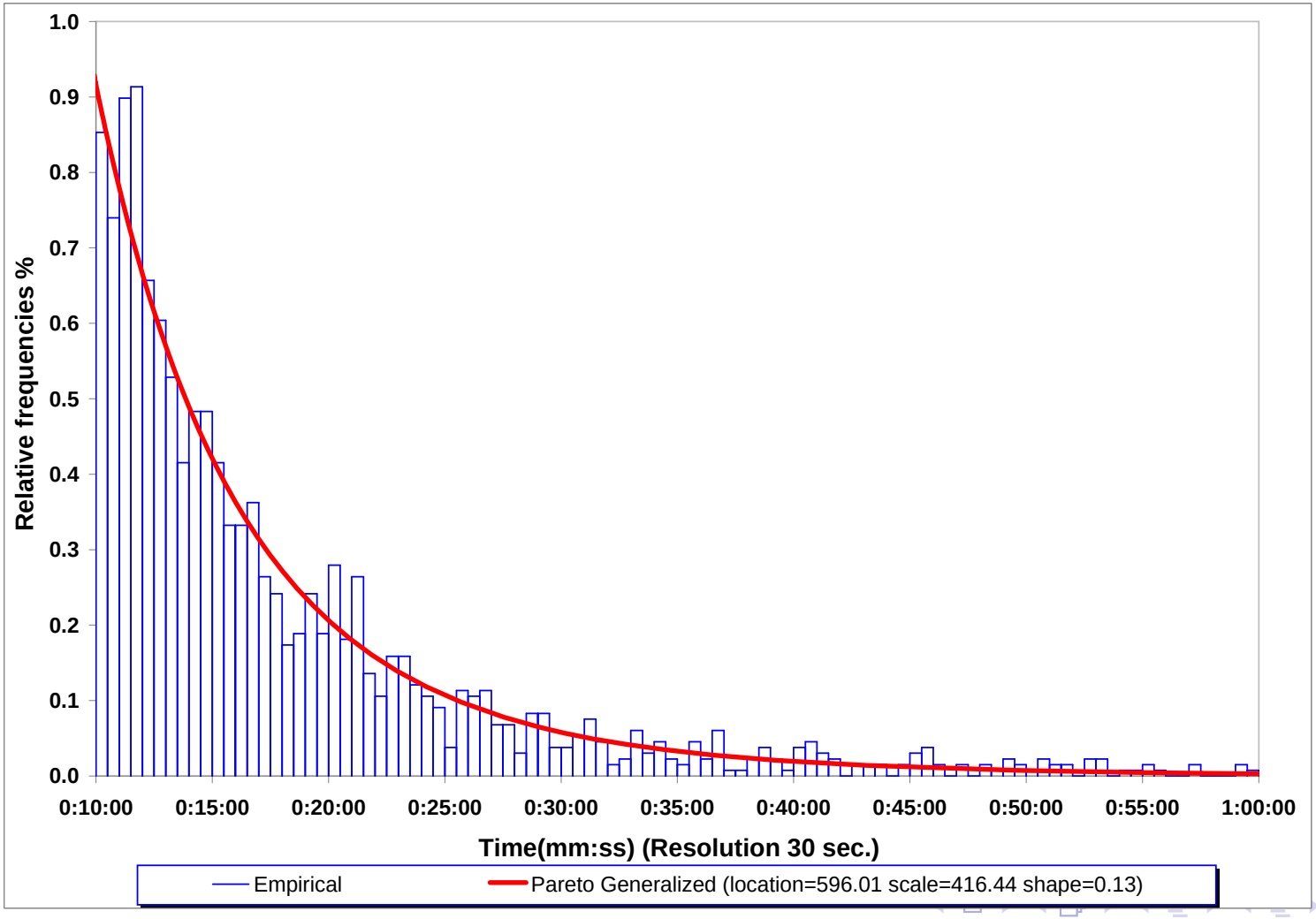
Patient Customers, Announcements, Priority Upgrades



Primitives: (Im)Patience

Israeli Bank: Uncensored 13,000 Customers, 24/11/2008

Patience $\geq 10min$: Why Pareto Tail?

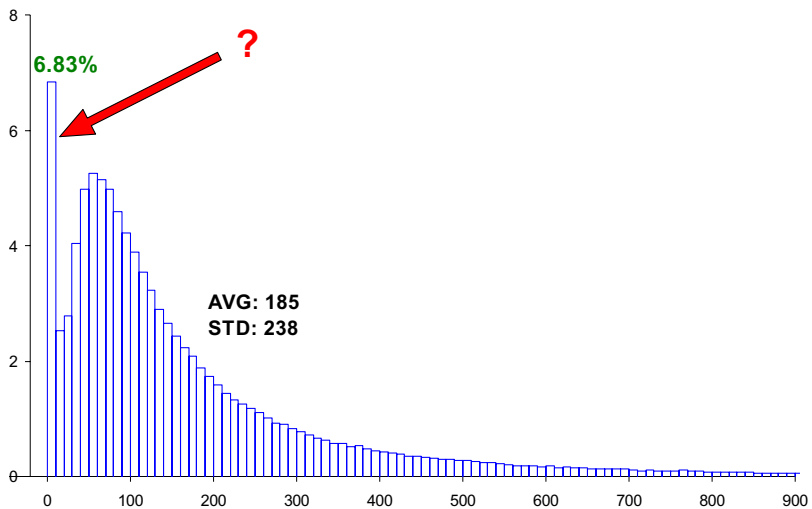


Primitives: Services (Durations)

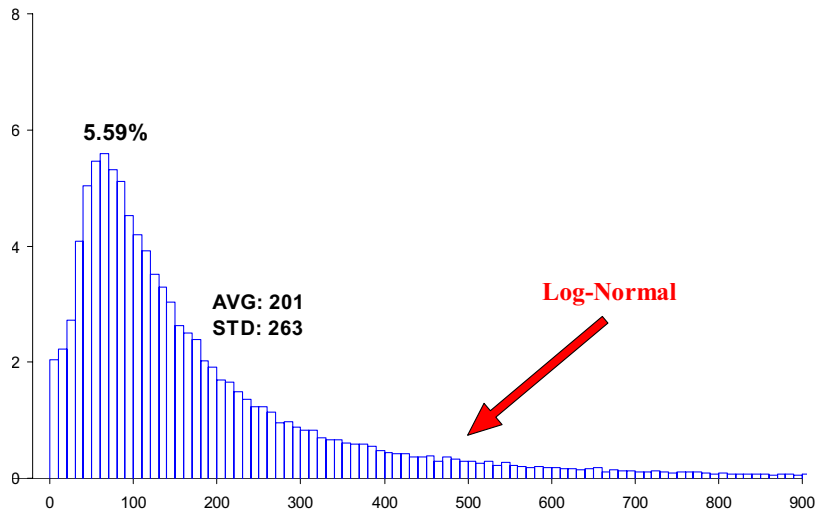
Histogram of Service-Duration in an Israeli Call Center, 1999

Why short services? Why LogNormal?

January-October



November-December



- January-October: **6.8% Short-Services** ($\leq$ 10 seconds) ?
- November-December: **LogNormal** durations (common) ?

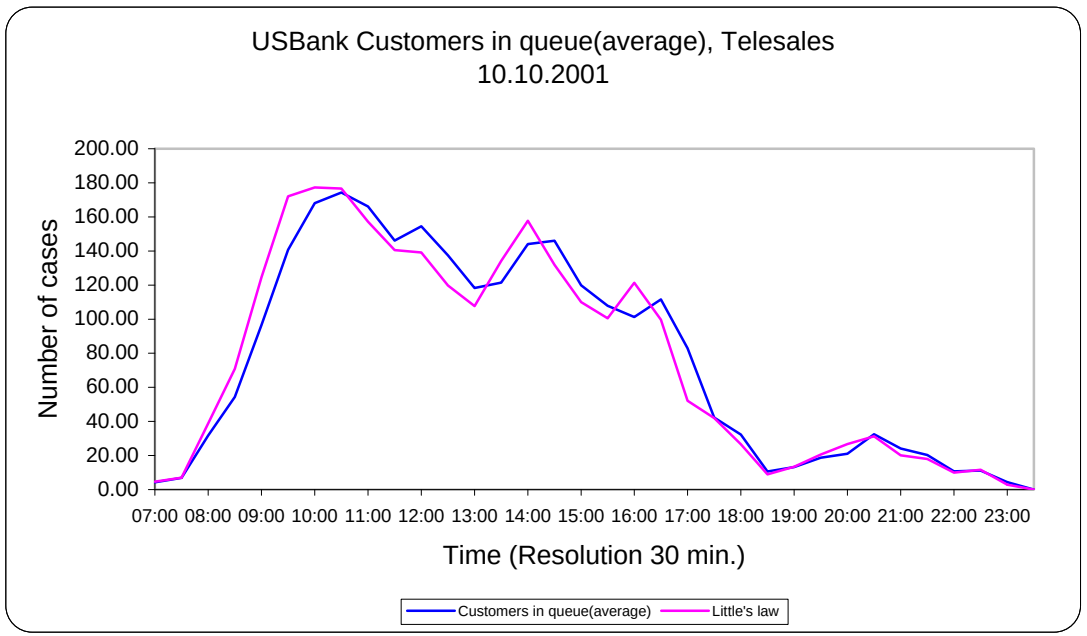
Conference Day Three: 11 Sep 2014 (Thursday)

Rooms: Before lunch: Blauwe Zaal (Auditorium)
After lunch: Kennispoort

8:00 – 9:00	Registrations
9:00 – 10:30	**Session 9: Process Analytics** *Session Chair:* Jörg Desel *Beyond Tasks and Gateways: Discovering BPMN Models with Subprocesses, Boundary Events and Activity Markers* Raffaele Conforti, Marlon Dumas, Luciano García-Bañuelos and Marcello La Rosa *Behavioral Comparison of Process Models Based on Canonically Reduced Event Structures* Abel Armas-Cervantes, Paolo Baldan, Marlon Dumas and Luciano García-Bañuelos *Where did I go wrong? - Explaining errors in business process models* Niels Lohmann and Dirk Fahland
10:30 – 11:00	Morning Tea / Coffee (Auditorium)
11:00 – 12:30	**Session 10: Resource and Time Management in BPM** *Session Chair:* Jan Mendling *Mining Resource-Scheduling Protocols* Arik Senderovich, Matthias Weidlich, Avigdor Gal and Avishai Mandelbaum *Dealing with Changes of Time-Aware Processes* Andreas Lanz and Manfred Reichert *Temporal Anomaly Detection in Business Processes* Andreas Rogge-Solti and Gjergji Kasneci

Little's Law $L = \lambda \times W$, in a Time-Varying Environment

Time-Gap: **# in System** lags behind **Little / 30 min**

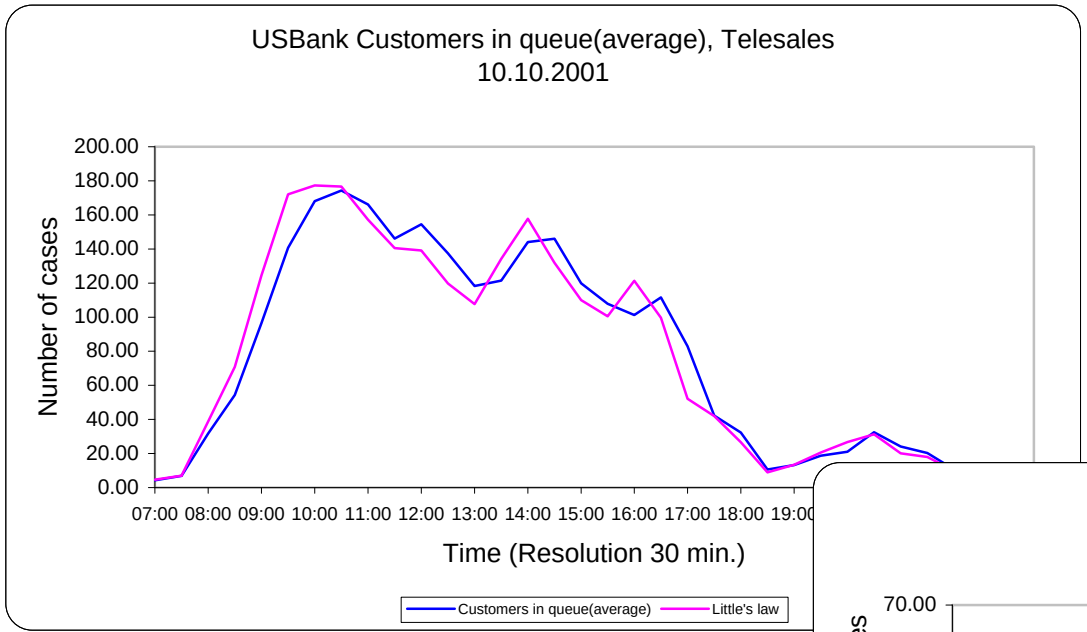


Call Center:

$\Rightarrow$ **Piecewise Steady-State**

Little's Law $L = \lambda \times W$, in a Time-Varying Environment

Time-Gap: **# in System** lags behind **Little / 30 min**



Call Center:

$\Rightarrow$ **Piecewise Steady-State**

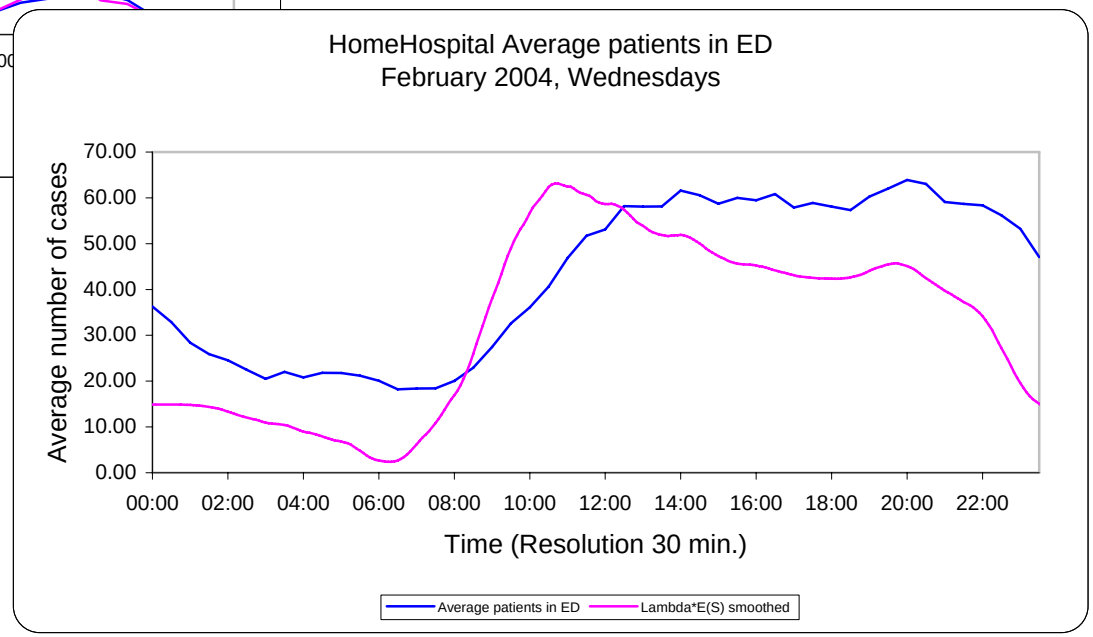
Emergency Dept:

$\Rightarrow$ **Time-Varying Transient**

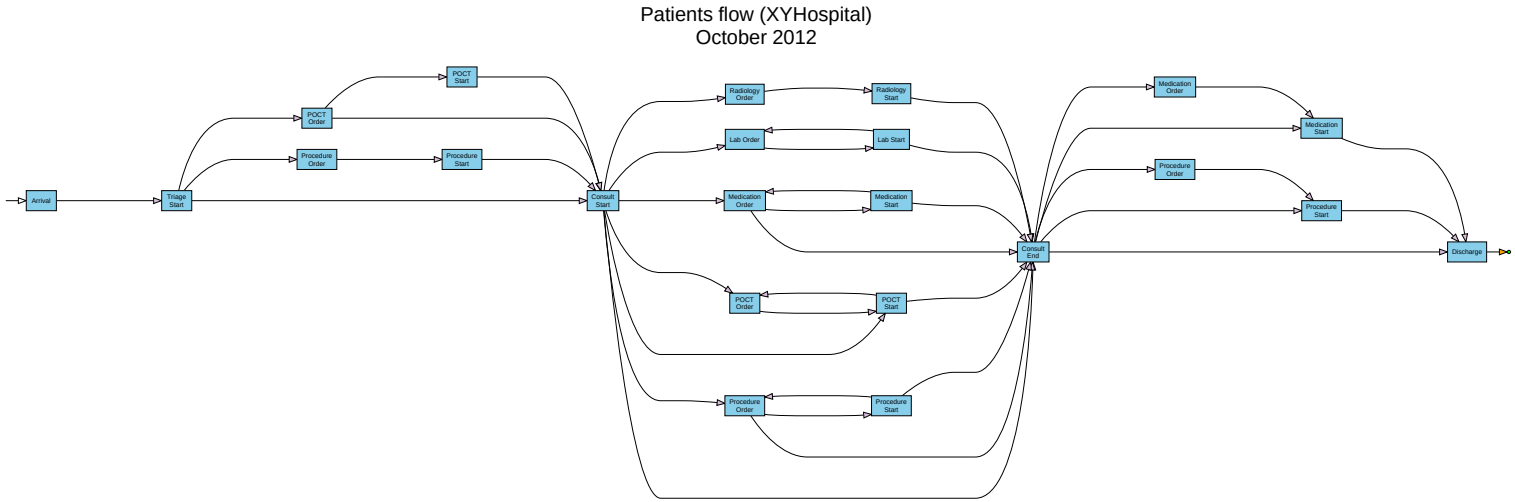
$$EL(t) = \tilde{\lambda}(t) \times EW,$$

$$\tilde{\lambda}(t) = E\lambda(t - W_e).$$

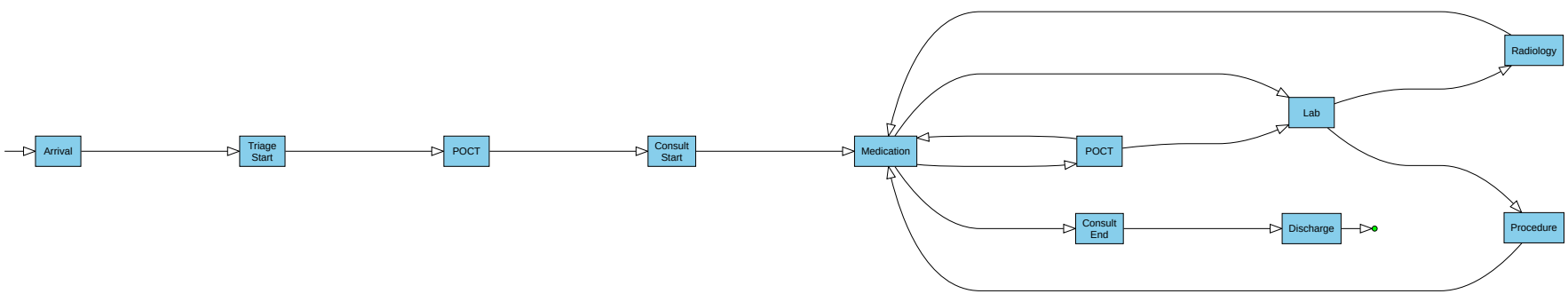
(Bertsimas, Mourtzinou; Fralix, Riano, Serfozo)



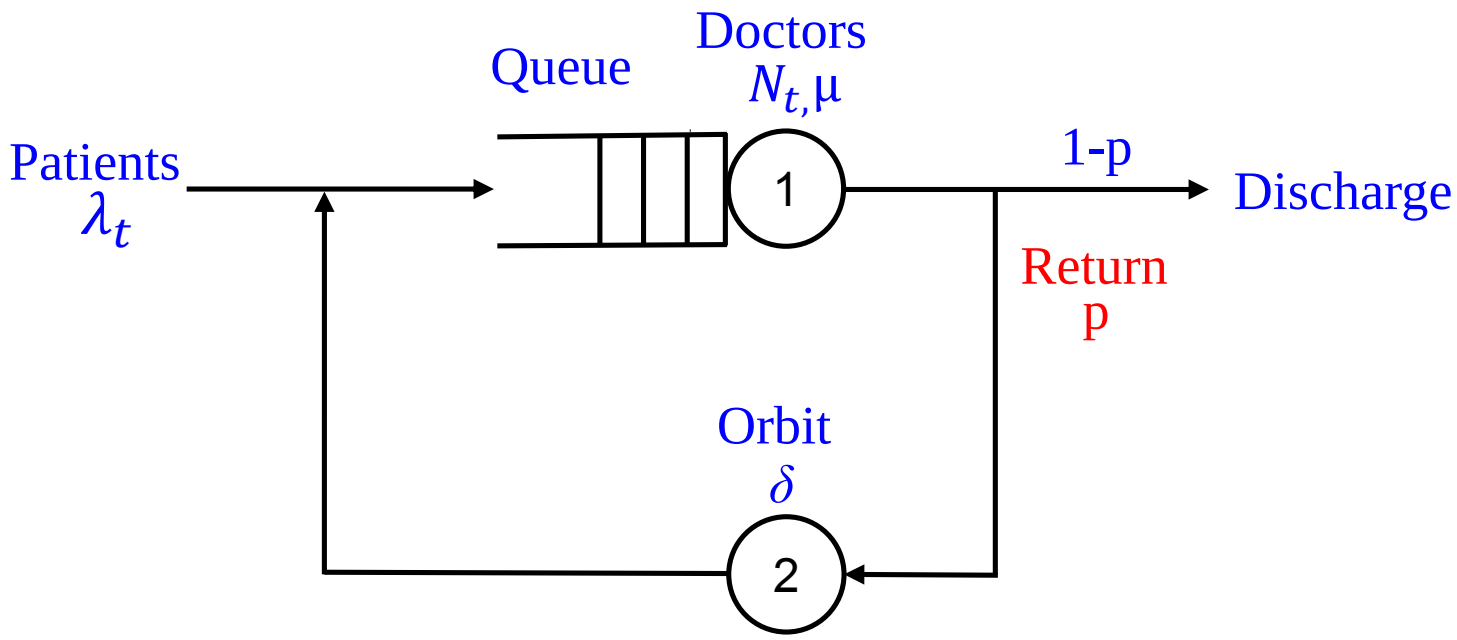
Recurrent Service Process in the ED



Capture **Recurrent** nature of service process: **Multiple doctor visits**

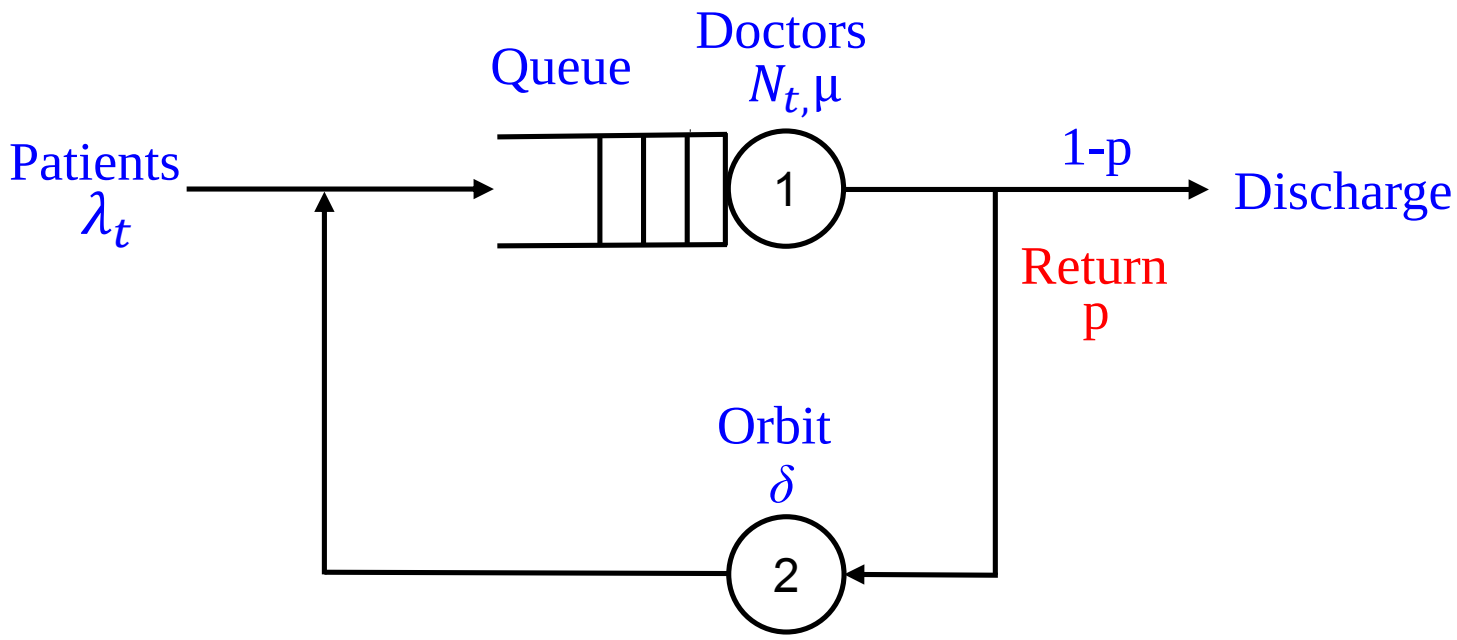


The Basic Service-Network Model: Erlang-R



w/ G. Yom-Tov

The Basic Service-Network Model: Erlang-R



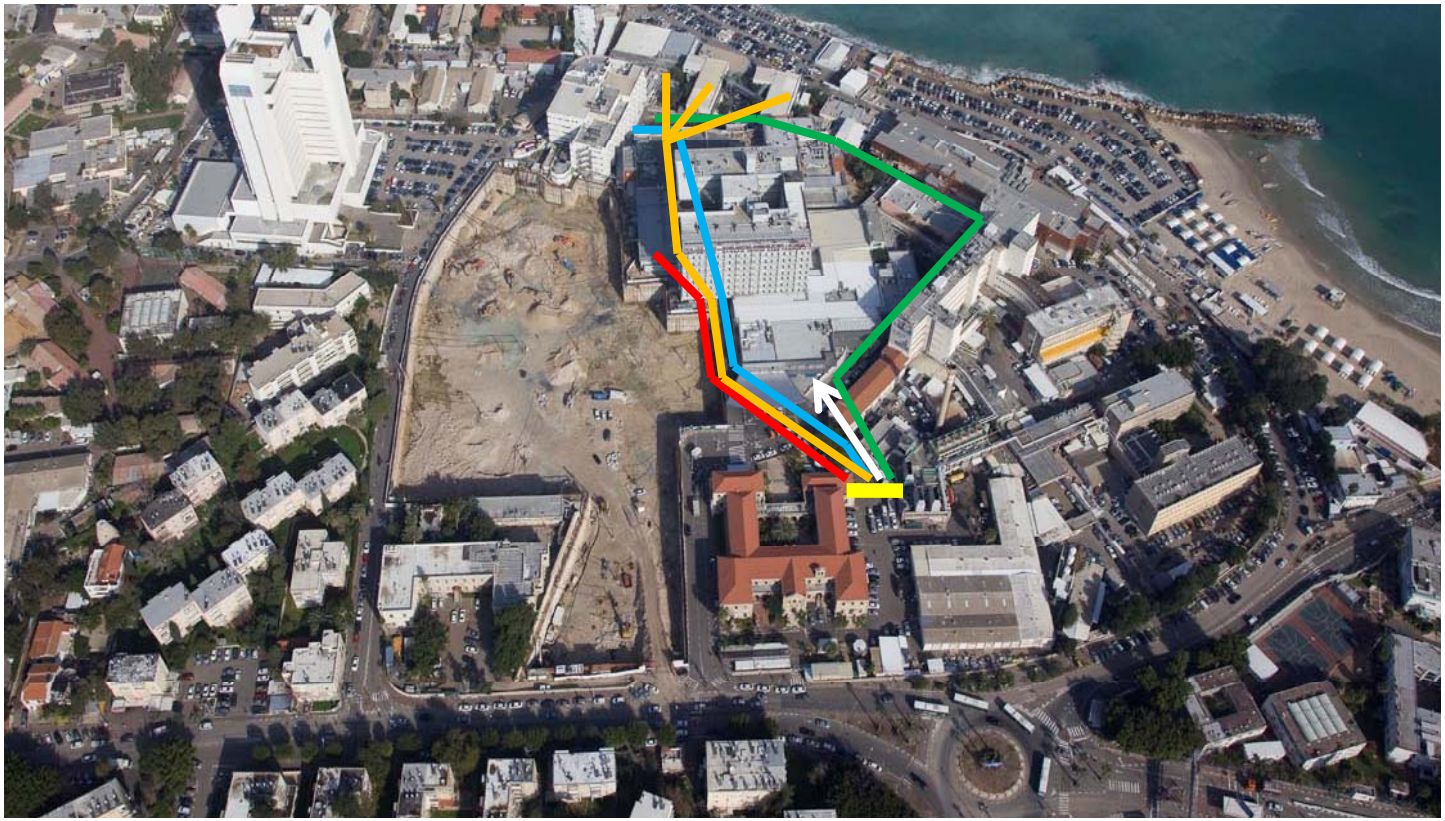
w/ G. Yom-Tov

2-station "Jackson" Network = (M/M/S, M/M/∞) :

- λ_t – **Time-Varying Arrival** rate
- N_t – Number of **Servers** (Physicians, or Nurses)
- μ – **Service** rate (E[Service] $= \frac{1}{\mu}$)
- p – **Return** (ReEntrant) fraction
- δ – **Orbit-to-Queue** rate (E[Delay]$_{20} = \frac{1}{\delta}$)

RFID-Based Data in Mass Casualty Event (Drill)

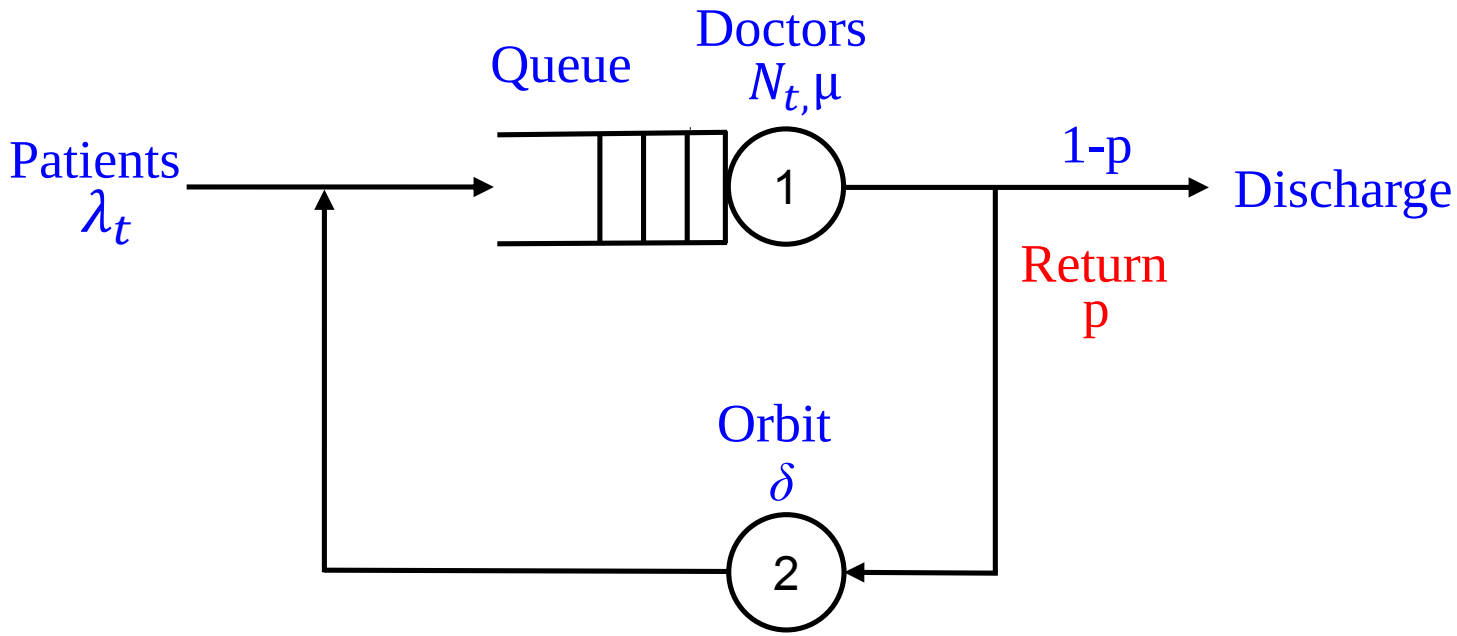
Chemical MCE, Rambam Hospital (May 2010, 11:00-13:00)



Fluid Model:

- **Predictable** Variability $\Rightarrow$ **Time-Varying**
- **Stochastic** Individualism averaged-out $\Rightarrow$ **Deterministic**

Fluid Model $\leftrightarrow$ (Time-Varying) Erlang-R System



Functional Strong Law of Large Numbers, for a 2-station QNet. BUT

FNet = ODE: derived **directly** (no QNet), spreadsheet "solution"

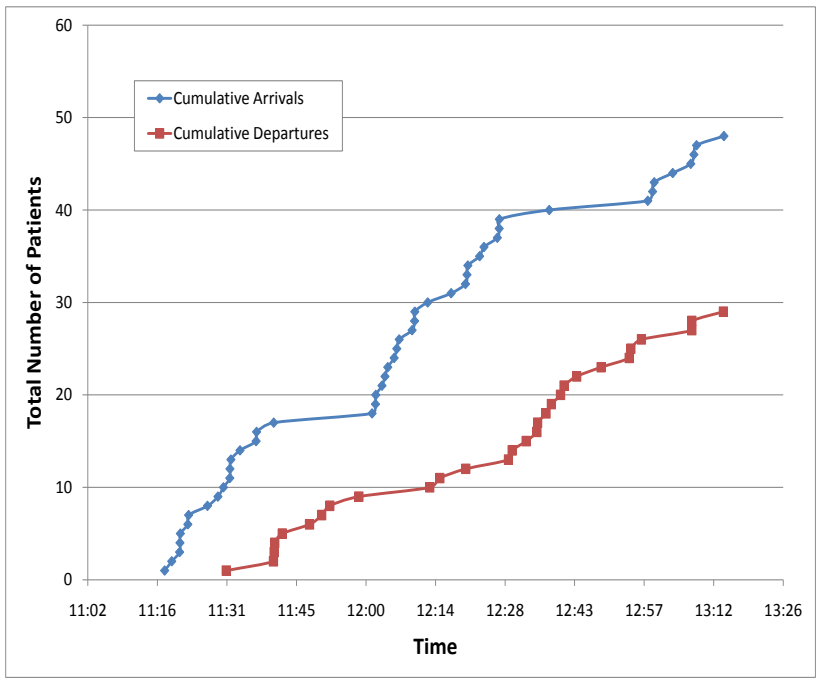
$$\frac{d}{dt} q_t^1 = \lambda_t - \mu \cdot (q_t^1 \wedge N_t) + \delta \cdot q_t^2$$

$$\frac{d}{dt} q_t^2 = p \cdot \mu \cdot (q_t^1 \wedge N_t) - \delta \cdot q_t^2$$

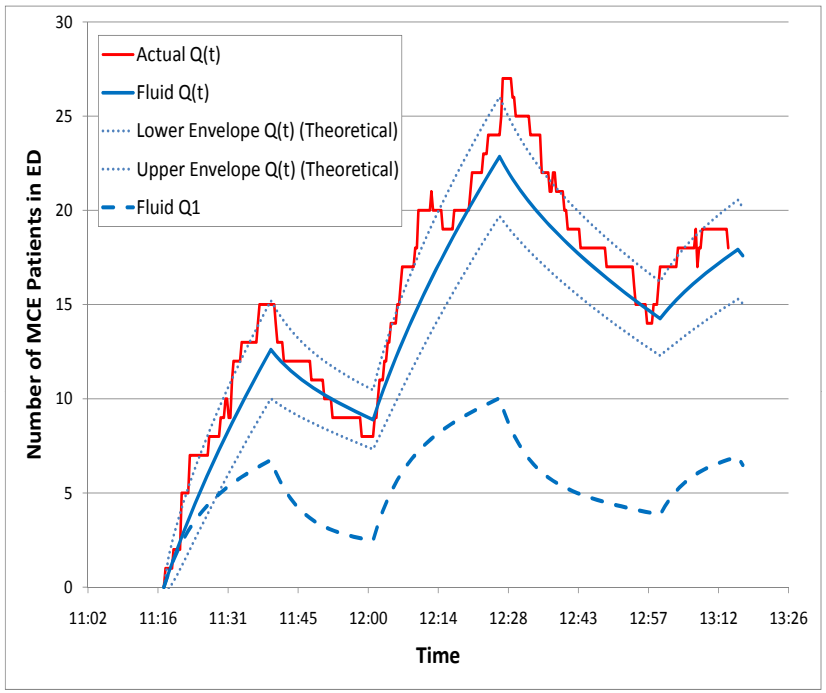
Erlang-R Value: FNet vs. Data

Chemical MCE Drill (Israel, May 2010, 11:00-13:00)

Arrivals & **Departures** (RFID)



Erlang-R (**Fluid** , **Diffusion**)

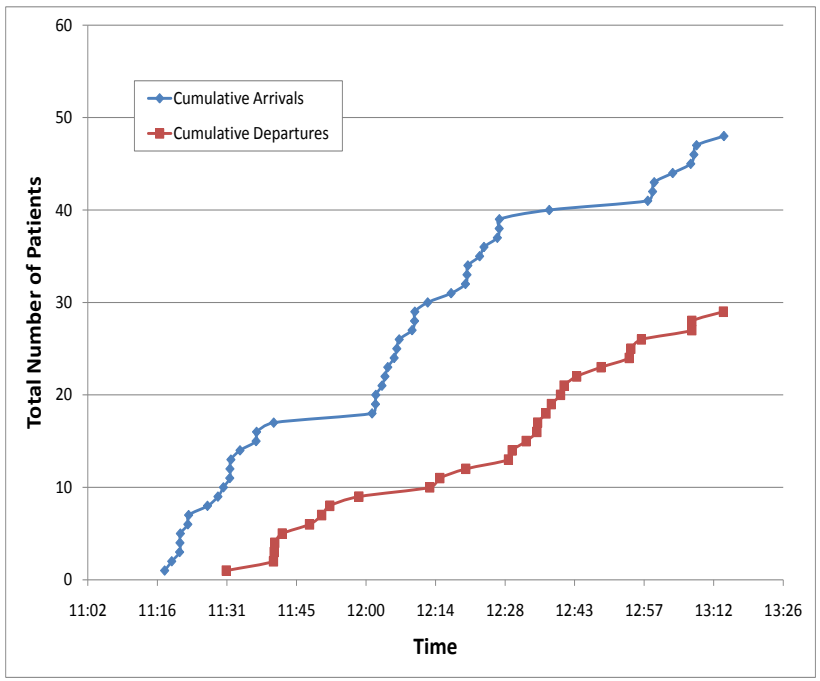


- **Recurrent/Repeated** services in Chemical MCE: injection every 15/30/60 min

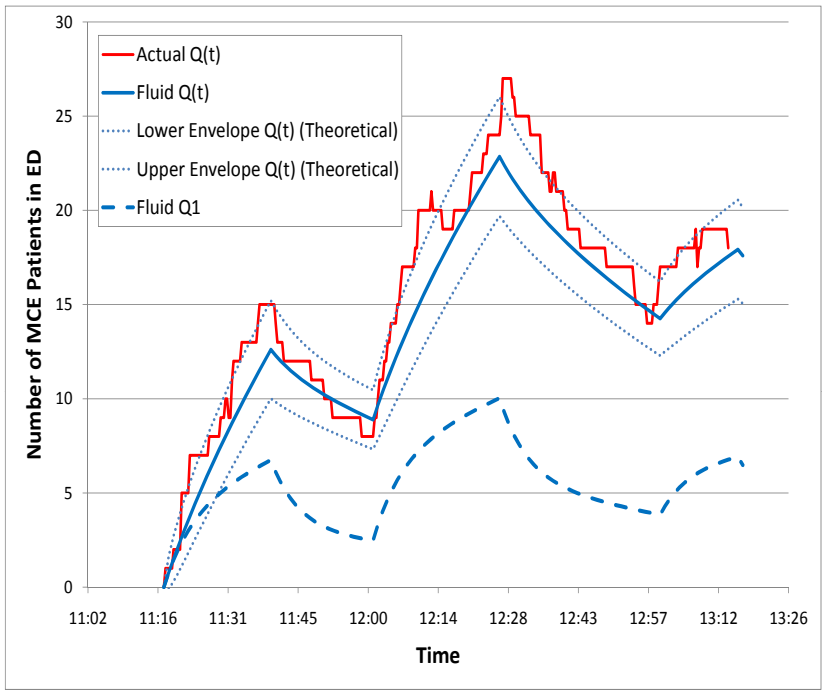
Erlang-R Value: FNet vs. Data

Chemical MCE Drill (Israel, May 2010, 11:00-13:00)

Arrivals & **Departures** (RFID)



Erlang-R (Fluid , Diffusion)



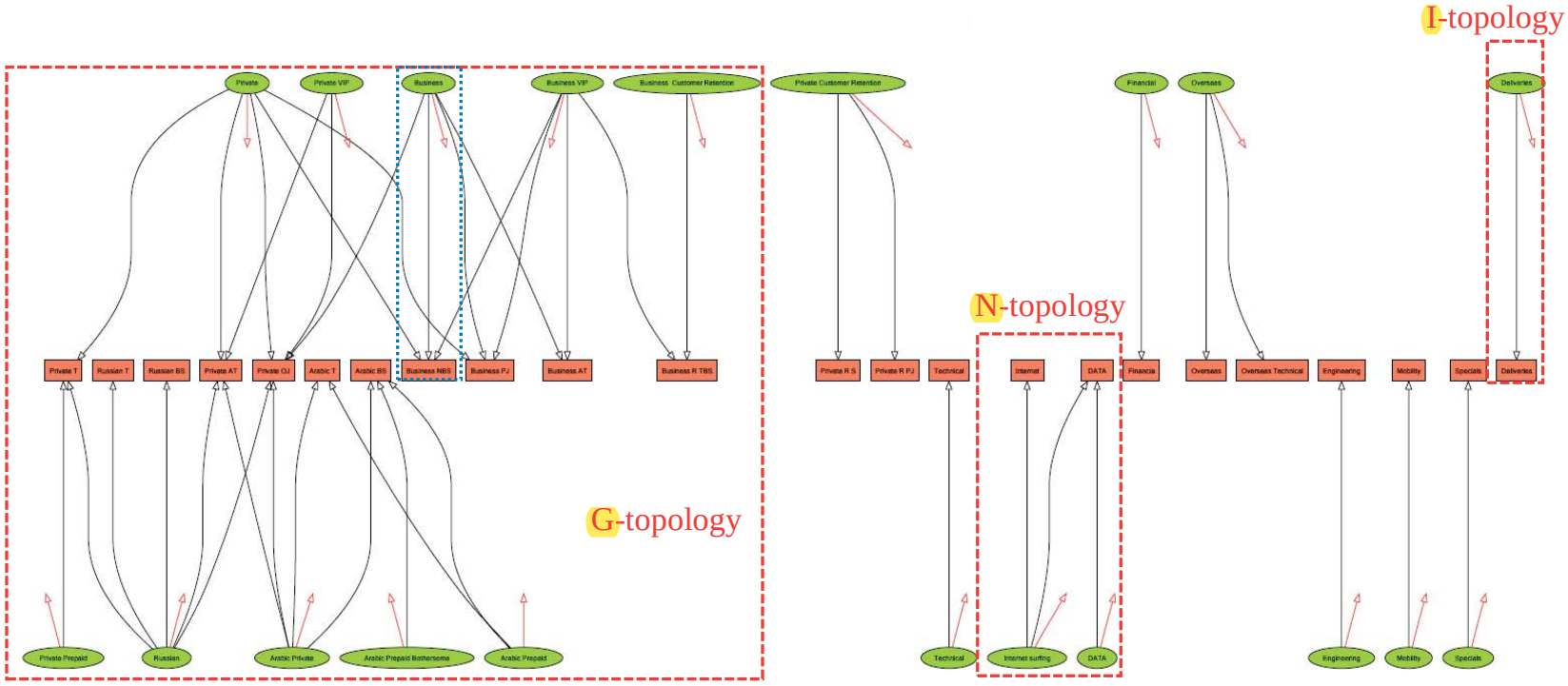
- **Recurrent/Repeated** services in Chemical MCE: injection every 15/30/60 min
- **Fluid** = ODE
- **Diffusion** (confidence band), via F. Central Limit Theorem: Usefully narrow

A Data-Based Framework, or “Erlang-R in the ED"

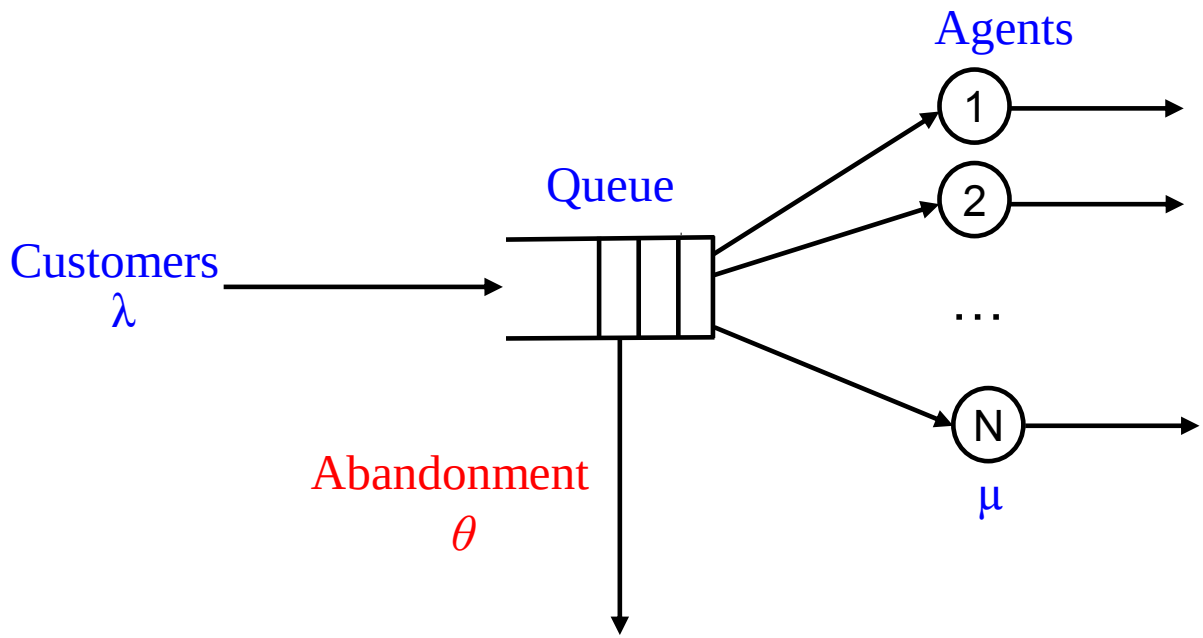
System = e.g. Emergency Department

- **QNet** = Erlang-R (time-varying 2-station Jackson; w/ Yom-Tov)
- **FNets** = 2-dim dynamical system (Massey & Whitt)
- **DNets** = 2-dim Markovian Service Net (w/ Massey and Reiman)
- **SimNet** = Customized ED-Simulator (Marmor & Sinreich)

ILTelecom 9/3/2008



A Basic Staffing Model: Erlang-A



w/ O. Garnett

"Birth & Death" Queue = M/M/N + M (Palm 1940's):

- λ – **Arrival** rate (Poisson)
- μ – **Service** rate (Exponential; $E[S] = \frac{1}{\mu}$)
- θ – **Patience** rate (Exponential, $E[\text{Patience}] = \frac{1}{\theta}$)
- N – Number of **Servers** (Agents).

Asymptotic Landscape: 9 Operational Regimes, and then some

Erlang-A, w/ I. Gurvich & J. Huang

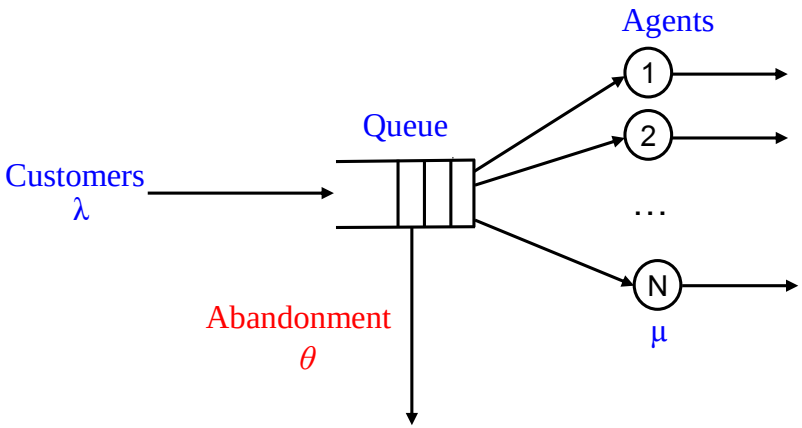
Erlang-A μ & θ fixed	Conventional scaling			Many-Server scaling			NDS scaling		
	Sub	Critical	Over	QD	QED	ED	Sub	Critical	Over
Offered load per server	$\frac{1}{1+\delta}$	$1-\frac{\beta}{\sqrt{n}}$	$\frac{1}{1-\gamma}$	$\frac{1}{1+\delta}$	$1-\frac{\beta}{\sqrt{n}}$	$\frac{1}{1-\gamma}$	$\frac{1}{1+\delta}$	$1-\frac{\beta}{n}$	$\frac{1}{1-\gamma}$
Arrival rate λ	$\frac{\mu}{1+\delta}$	$\mu-\frac{\beta}{\sqrt{n}}\mu$	$\frac{\mu}{1-\gamma}$	$\frac{n\mu}{1+\delta}$	$n\mu-\beta\mu\sqrt{n}$	$\frac{n\mu}{1-\gamma}$	$\frac{n\mu}{1+\delta}$	$n\mu-\beta\mu$	$\frac{n\mu}{1-\gamma}$
# servers	1			n			n		
Time-scale	n			1			n		
Impatience rate	θ/n			θ			θ/n		
Staffing level	$\frac{\lambda}{\mu}(1+\delta)$	$\frac{\lambda}{\mu}(1+\frac{\beta}{\sqrt{n}})$	$\frac{\lambda}{\mu}(1-\gamma)$	$\frac{\lambda}{\mu}(1+\delta)$	$\frac{\lambda}{\mu}+\beta\sqrt{\frac{\lambda}{\mu}}$	$\frac{\lambda}{\mu}(1-\gamma)$	$\frac{\lambda}{\mu}(1+\delta)$	$\frac{\lambda}{\mu}+\beta$	$\frac{\lambda}{\mu}(1-\gamma)$
Utilization	$\frac{1}{1+\delta}$	$1-\sqrt{\frac{\theta}{\mu}}\frac{h(\hat{\beta})}{\sqrt{n}}$	1	$\frac{1}{1+\delta}$	$1-\sqrt{\frac{\theta}{\mu}}\frac{\hat{h}(\hat{\beta})}{\sqrt{n}}$	1	$\frac{1}{1+\delta}$	$1-\sqrt{\frac{\theta}{\mu}}\frac{h(\hat{\beta})}{n}$	1
$\mathbb{E}(Q)$	$\frac{1}{\delta(1+\delta)}$	$\sqrt{n}g(\hat{\beta})$	$\frac{n\mu\gamma}{\theta(1-\gamma)}$	$\frac{1}{\delta}\varrho_n$	$\sqrt{n}g(\hat{\beta})\alpha$	$\frac{n\mu\gamma}{\theta(1-\gamma)}$	$o(1)$	$ng(\hat{\beta})$	$\frac{n^2\mu\gamma}{\theta(1-\gamma)}$
$\mathbb{P}(Ab)$	$\frac{1}{n}\frac{1}{\delta}\frac{\theta}{\mu}$	$\frac{\theta}{\sqrt{n}\mu}g(\hat{\beta})$	γ	$\frac{1}{n}\frac{(1+\delta)}{\delta}\frac{\theta}{\mu}\varrho_n$	$\frac{\theta}{\sqrt{n}\mu}g(\hat{\beta})\alpha$	γ	$o(\frac{1}{n^2})$	$\frac{\theta}{n\mu}g(\hat{\beta})$	γ
$\mathbb{P}(W_q>0)$	$\frac{1}{1+\delta}$	≈ 1		ϱ_n	$\alpha\in(0,1)$	≈ 1	≈ 0	≈ 1	
$\mathbb{P}(W_q>T)$	$\frac{1}{1+\delta}e^{-\frac{\delta}{1+\delta}\mu T}$	$1+O(\frac{1}{\sqrt{n}})$	$1+O(\frac{1}{n})$	≈ 0		$f(T)$	≈ 0	$\frac{\bar{\Phi}(\hat{\beta}+\sqrt{\theta\mu}T)}{\bar{\Phi}(\hat{\beta})}$	$1+O(\frac{1}{n})$
Congestion $\frac{\mathbb{E}W_q}{\mathbb{E}S}$	$\frac{1}{\delta}$	$\sqrt{n}g(\hat{\beta})$	$n\mu\gamma/\theta$	$\frac{1}{n}\frac{(1+\delta)}{\delta}\varrho_n$	$\frac{\alpha}{\sqrt{n}}g(\hat{\beta})$	$\frac{\mu\gamma}{\theta}$	$o(\frac{1}{n})$	$g(\hat{\beta})$	$n\mu\gamma/\theta$

- Conventional: Ward & Glynn (03, $G/G/1+G$)
- Many-Server:
  - QED: Halfin-Whitt (81), w/ Garnett & Reiman (02)
  - ED: Whitt (04)
  - NDS: Atar (12)
- **"Missing"**: ED+QED; Hazard-rate scaling (M/M/N+G); Time-Varying, Non-Parametric; Moderate- and Large-Deviation; Networks (multi-regimes)

Asymptotic Erlang-X (Markovian Q's)

- Pre-History, 1914: **Erlang** (Erlang-B = M/M/n/n, Erlang-C = M/M/n)
- Pre-History, 1974: Jagerman (Erlang-B)
- History Milestone, 1981: **Halfin-Whitt** (Erlang-C, GI/M/n)
- Erlang-A (M/M/N+M), 2002: w/ **Garnett** & Reiman
- Erlang-A with General (Im)Patience (M/M/N+G), 2005: w/ Zeltyn
- Erlang-C (ED+QED), 2009: w/ Zeltyn
- Erlang-B with Retrial, 2010(3): Avram, Janssen, van Leeuwaarden
- Refined Asymptotics (Erlang A/B/C, . . .), 2008-2013: Janssen, van Leeuwaarden, Zhang, Zwart
- Production Q's, 2011: Reed & Zhang
- Universal Erlang-A: w/ Gurvich & Huang
- Queueing Networks:
  - (Semi-)Closed: Nurse Staffing (Jennings & de Vericourt), CCs with IVR (w/ Khudiakov), **Erlang-R (w/ Yom-Tov), Erlang-S (w/ Azriel and Feigin)**
  - CCs with Abandonment and Retrials: w. Massey, Reiman, Rider, Stolyar
  - Markovian Service Networks: w/ Massey & Reiman
- Leaving out:
  - **Non-Exponential Service Times**: M/D/n (Erlang-D), G/Ph/n, · · · , G/GI/n+GI, Measure-Valued Diffusions
  - **Dimensioning** (Staffing): M/M/n, · · · , time-varying Q's, V- and Reversed-V, · · ·
  - **Control**: V-network, Reversed-V, · · · , SBRNets

Universal Approximations: Erlang-A (M/M/N + M)



w/ **I. Gurvich & J. Huang**

- **QNet**: Birth & Death Queue, with B - D rates

$$F(q) = \lambda - \mu \cdot (q \wedge n) - \theta \cdot (q - n)^+, \quad q = 0, 1, \ldots$$

- **FNet**: Dynamical (Deterministic) System – ODE

$$dx_t = F(x_t)dt, \quad t \geq 0$$

- **DNet**: Universal (Stochastic) Approximation – SDE

$$dY_t = F(Y_t)dt + \sqrt{2\lambda}\, dB_t, \quad t \geq 0$$

Reconciling Time-Varying and Steady-State Models

- **Rigid** (fixed) staffing level during a time-varying shift:
  Doomed to alternate between overloading and underloading
- **Flexible** staffing:
  Can design **time-varying staffing** that achieves, **at all times, Steady-State performance**
  via Square-Root Staffing (Modified Offered-Load)

- **History**:
  - Jennings, M., Reiman, Whitt (1996): Emergence of the phenomenon, with infinite-server heuristics
  - Feldman, M., Massey, Whitt (2008): Stabilize delay probability with QED staffing, with little theory
  - Liu and Whitt (2012): Stabilize abandonment probability, with ED theory
  - w/ Huang, Gurvich (ongoing): QED theory

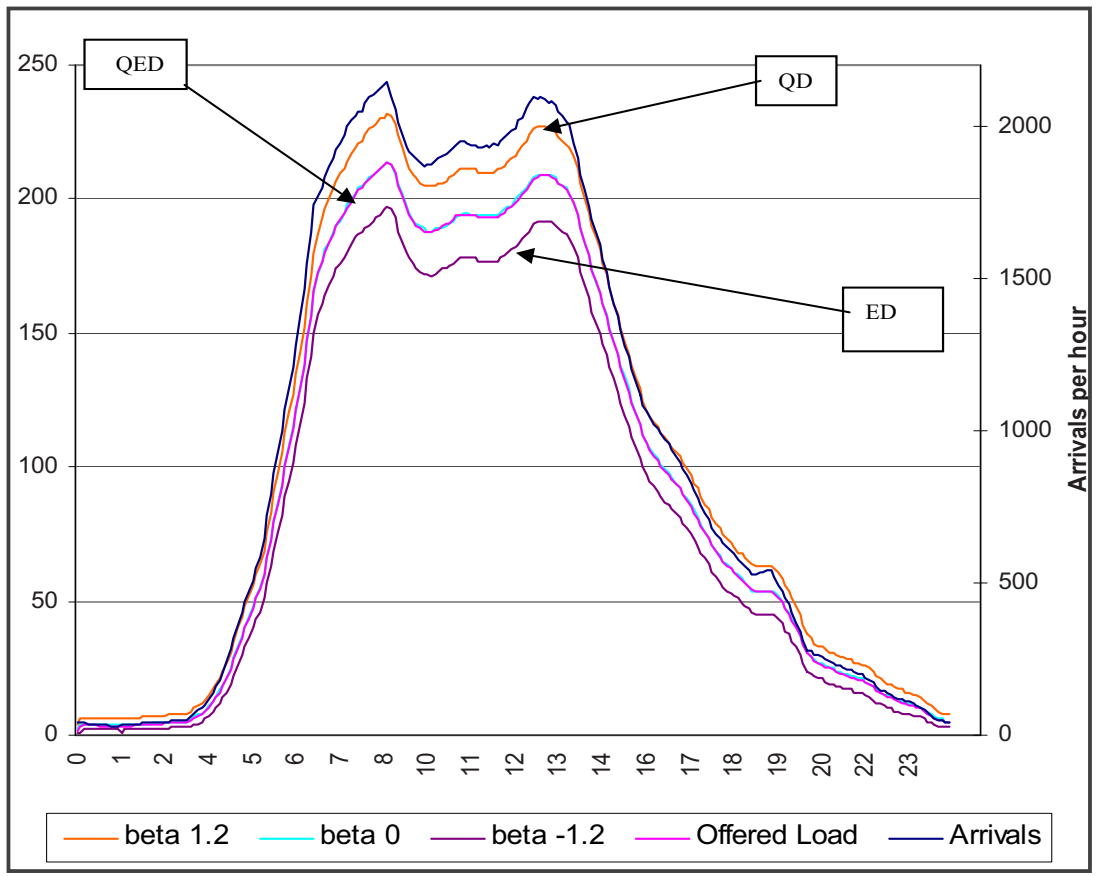
Time-Varying Arrival Rates

Square-Root Staffing:

$$N(t) = R(t) + \beta\sqrt{R(t)}, \quad -\infty < \beta < \infty.$$

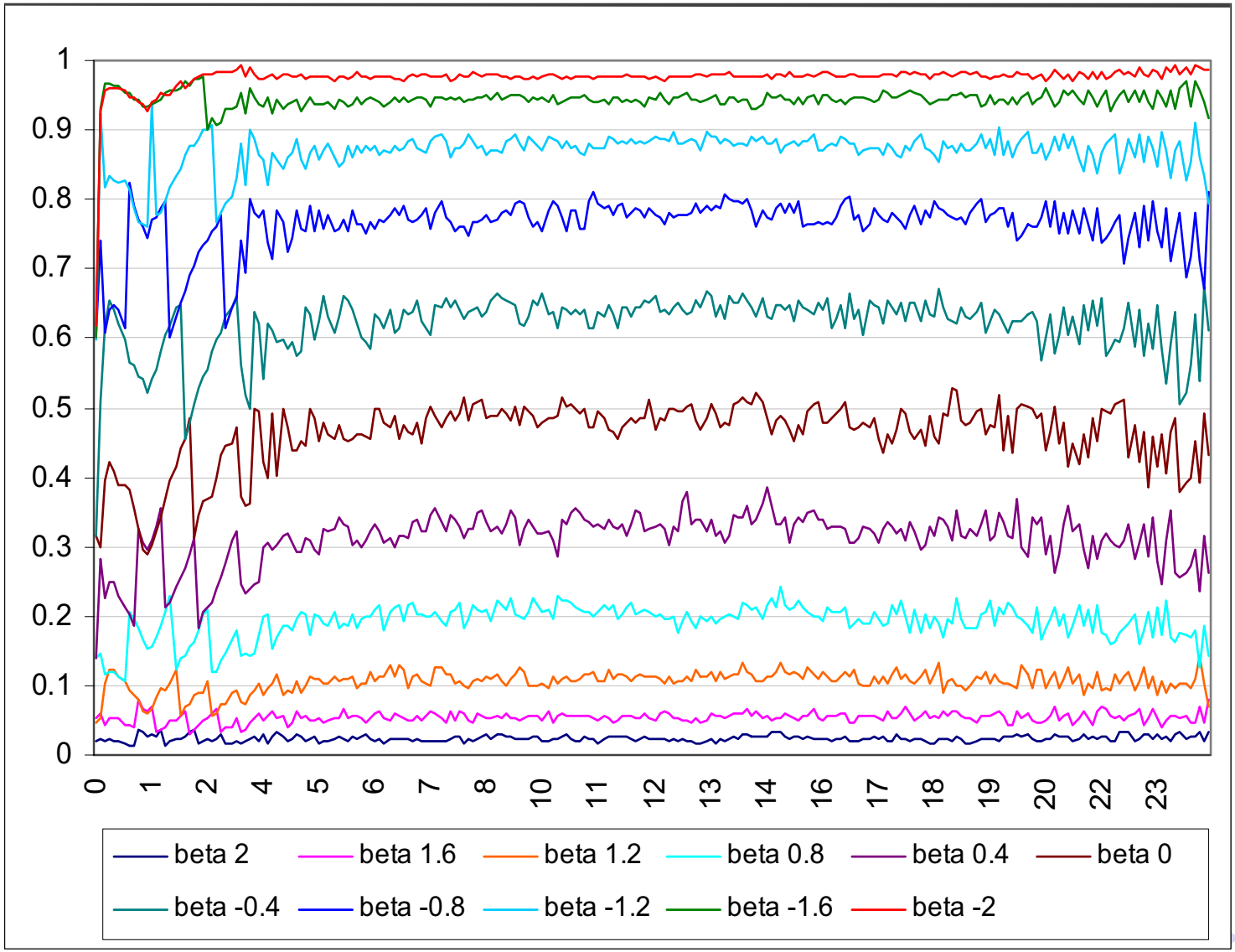
$R(t)$ is the **Offered-Load** at time t $\quad (R(t) \neq \lambda(t) \times \mathrm{E[S]})$

Arrivals, Offered-Load and Staffing



Time-Stable Performance of Time-Varying Systems

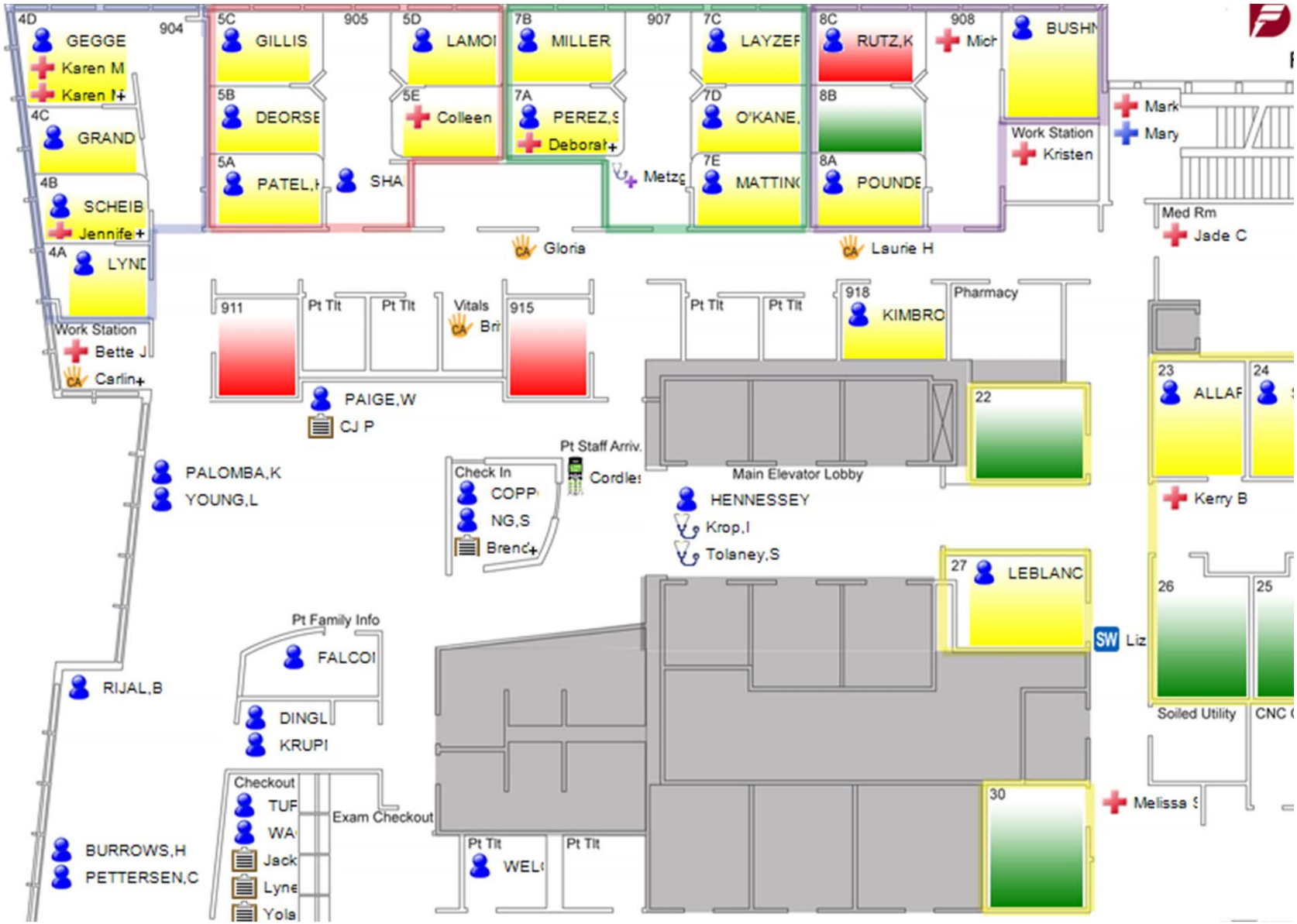
Delay Probability = as in the **Stationary Erlang-A / R**



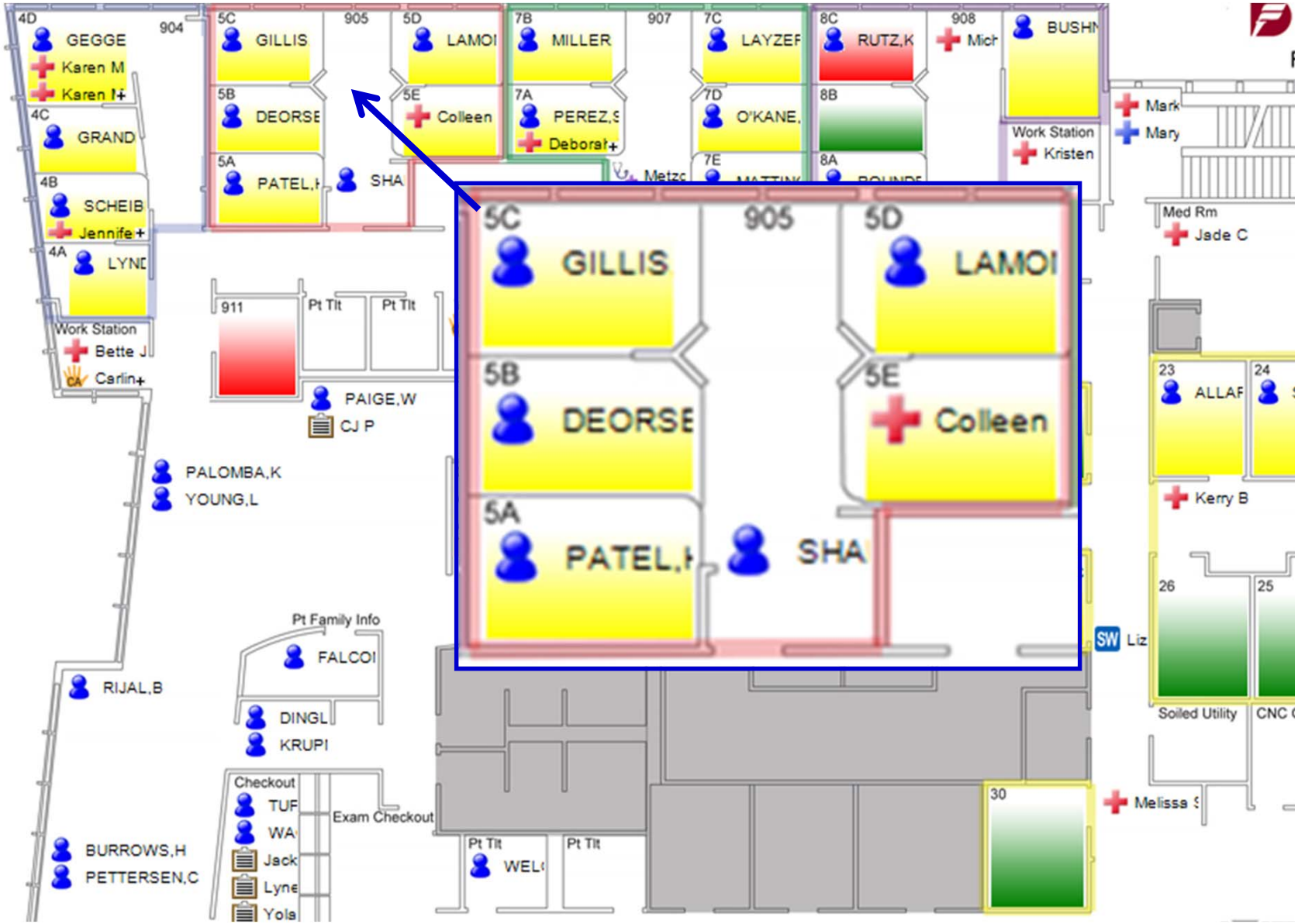
Data $\Leftrightarrow$ Questions, Directions, Territories

- **Open Questions**
  - Why are service durations **Log-Normal**?
  - Staffing to **stabilize** time-varying performance: w/ Huang & Gurvich
  - **Fork-join** networks (Petri Nets): w/ Atar & Zviran; Zaeid & Kaspi
- **New Directions**
  - Multi-channel **design**:
    self-service (eg. IVR, w/ Carmeli & Kaspi), tele-medicine
  - **Protocol** Inference: w/ Liberman & Meilijson; w/ Senderovic, Weidlich, Gal
  - **Personalized** Queueing Theory: w/ Momcilovic (LPF)
  - Offered-Load/Capacity calculations:
    operational, cognitive, **emotional; financial**
  - All models “born" **equal**:
    eg. inference via diffusions (w/ Pang), simulating FNets
- **Uncharted Territories**
  - **Server Nets**: w/ Armony & Momcilovic; Azriel & Feigin; Senderovic & Gal
  - **RTLS**-based research (OR/Stat/IS, e.g. real-time control)
  - (Semi)Automatic **creation** of SNets (OR + Process Mining)
  - **Integrated** healthcare:
    from onset of symptoms through the hospital to recovery

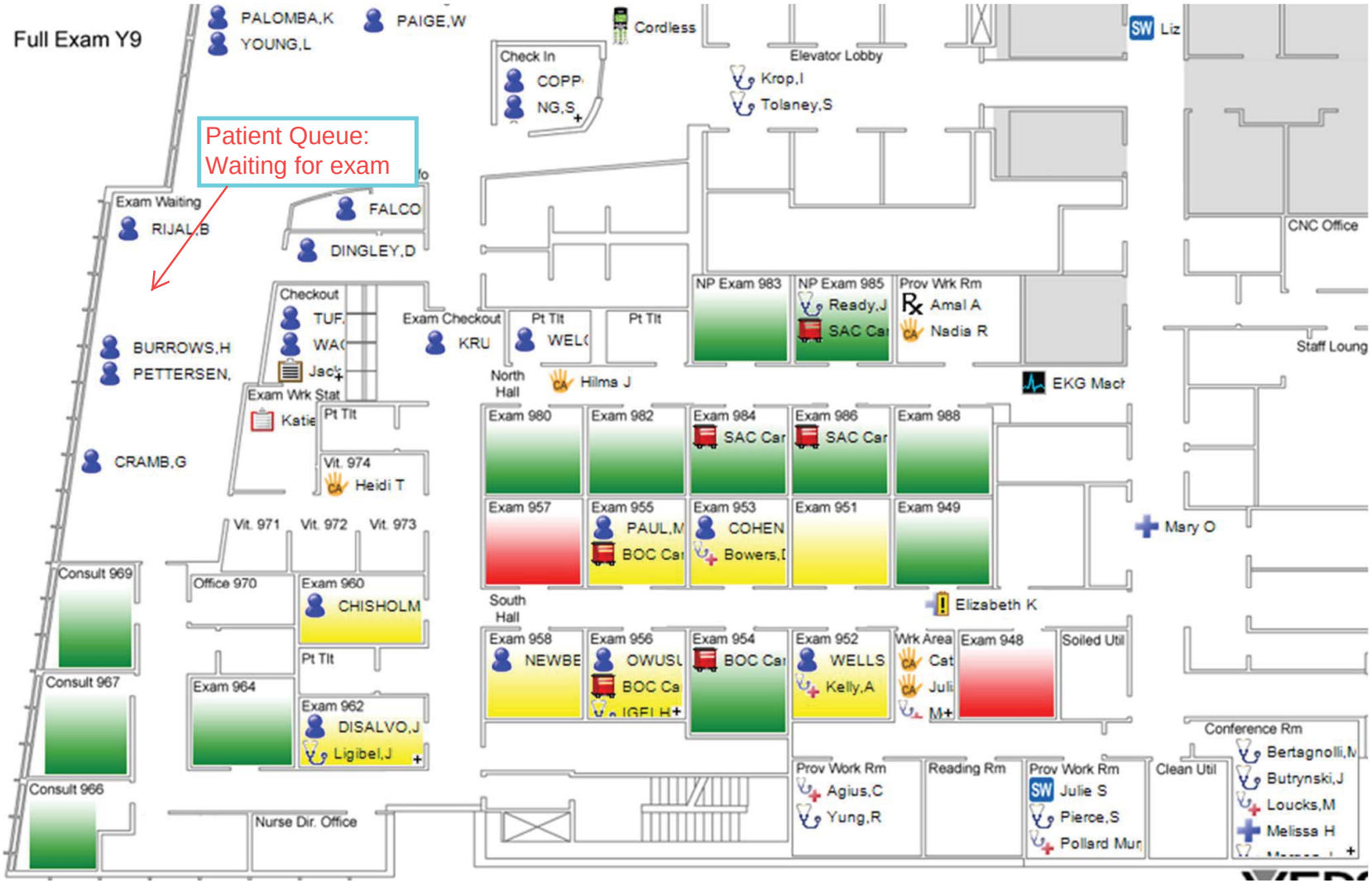
Uncharted: Closing the Gap via RTLS (1/2 9th Floor)



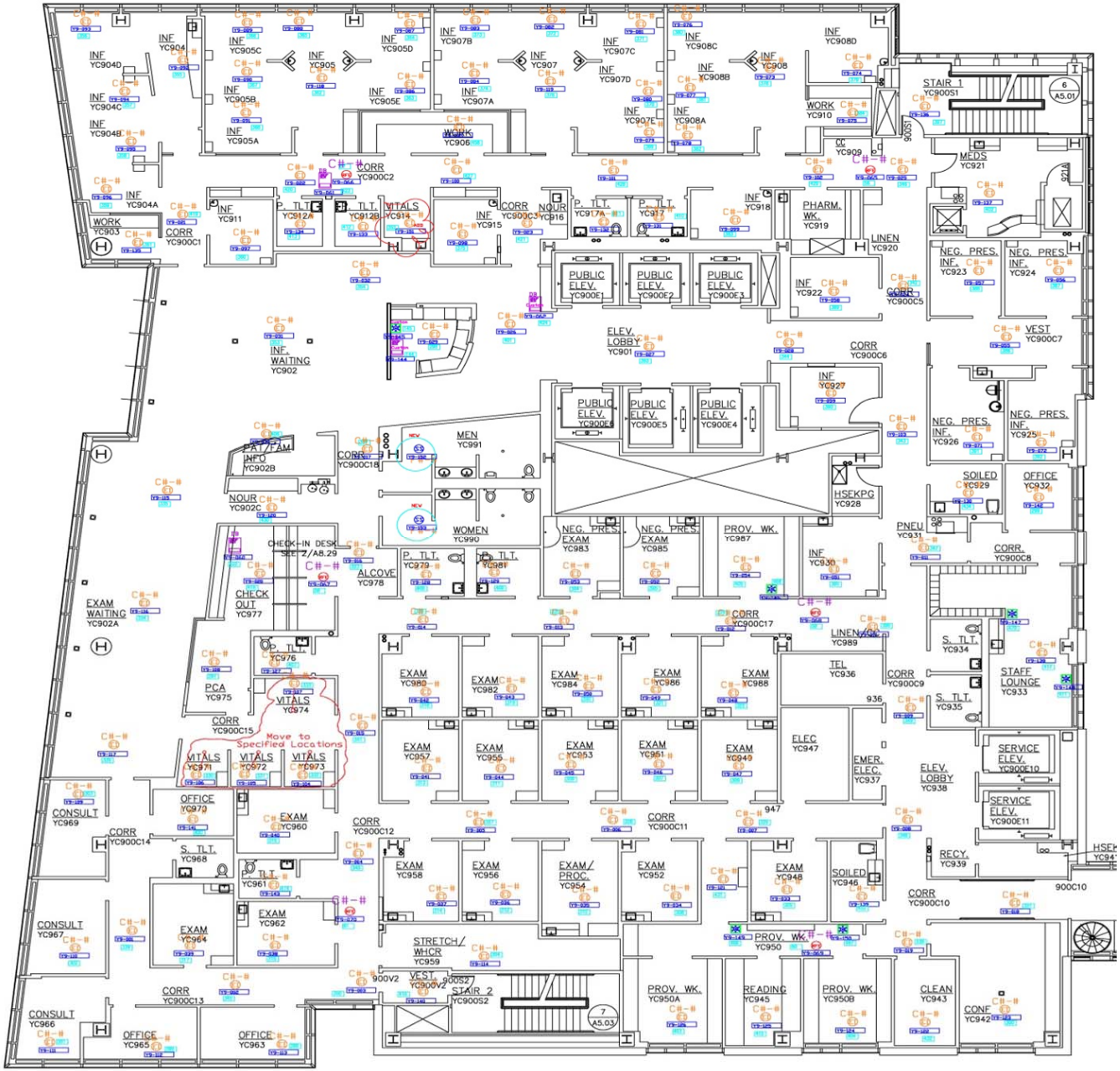
5 Patients in Room 905



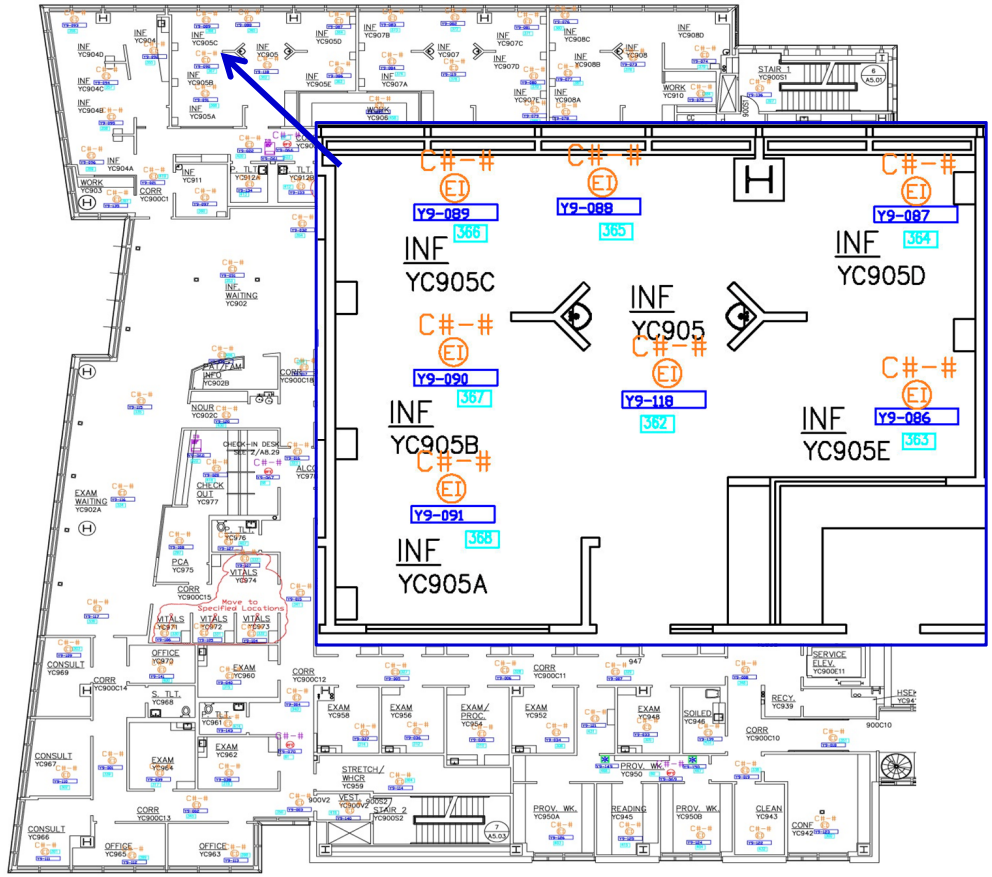
Queues, Exams, Conference (Second 1/2 of 9th Floor)



900 RTLS Readers, 3 seconds Resolution



7 RTLS Readers in Room 905, out of 900 Total



- **Data**: **3 sec** resolution, **250,000** events/day, covering:
  - **1000** patients per day
  - **400** staff
  - Equipment

RTLS Data Challenges: Validation and Processing at the SEELab

- Many months of data-processing, still ongoing: e.g. 6000 lines of code
  - Led by **Trofimov, with Gavako, Nadjharov, Senderovic**
  - Extremely generous support of **partner hospital !!!**
- 2 data sources - **Appointment Schedules** and **RTLS Receivers**: focus on latter but also use former

RTLS Data Challenges: Validation and Processing at the SEELab

- Many months of data-processing, still ongoing: e.g. 6000 lines of code
  - Led by **Trofimov, with Gavako, Nadjharov, Senderovic**
  - Extremely generous support of **partner hospital !!!**
- 2 data sources - **Appointment Schedules** and **RTLS Receivers**: focus on latter but also use former
- Validation process identified **noise** and **inconsistencies**
  - Denoised best we could
  - Corrected inconsistencies only when proved as such
  - **All documented**

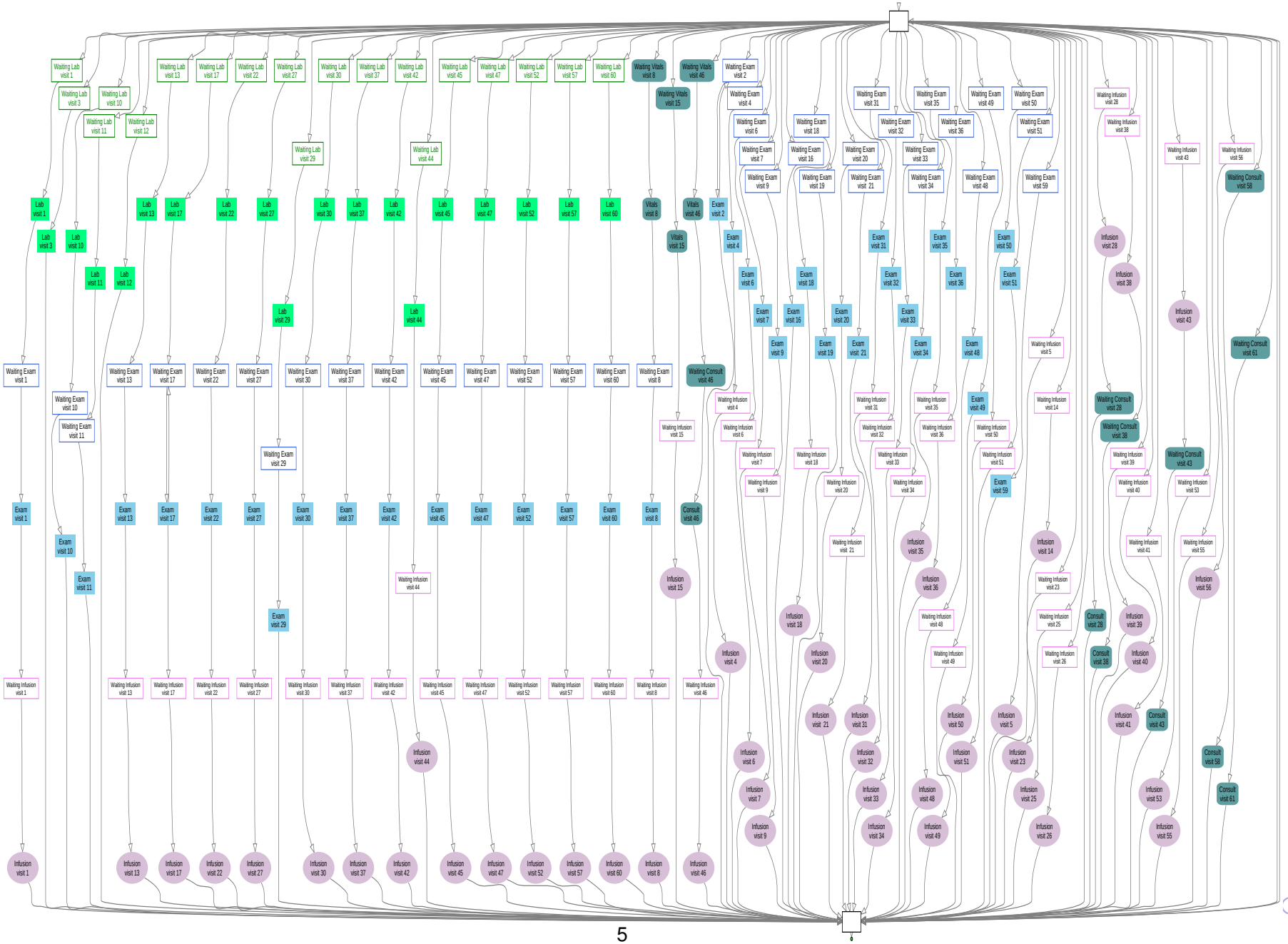
RTLS Data Challenges: Validation and Processing at the SEELab

- Many months of data-processing, still ongoing: e.g. 6000 lines of code
  - Led by **Trofimov, with Gavako, Nadjharov, Senderovic**
  - Extremely generous support of **partner hospital !!!**
- 2 data sources - **Appointment Schedules** and **RTLS Receivers**: focus on latter but also use former
- Validation process identified **noise** and **inconsistencies**
  - Denoised best we could
  - Corrected inconsistencies only when proved as such
  - **All documented**
- Gave rise to
  - **"Manual"**: Data description and documentation (77 pages)
  - **Exploratory Data Analysis**: Statistics and Animation (222 pages)

RTLS Data Challenges: Generic, Specific

- **Software errors**: e.g. negative duration values, few-second visits in rooms, floor-to-floor-to-back in seconds, . . .
- **Noncompliance** of patients and medical staff: e.g leaving badges in waiting areas, exam rooms or hallways; out of 1000 exams per day, 100 to rooms “without" physician - possibly fixable via appointments
- **Noisy** by nature: e.g few-second room visits, removed via threshold-denoising
- **Translating** locations to activities: e.g. avoiding fictitious visits
- **Matching** accurately patients with their service providers: e.g. patient arriving to exam room before MD, then examined, then leave after MD $\Rightarrow$ must read both patient and MD to calculate waiting & exam duration; could arise when a provider attends to several patients simultaneously
- **Planned vs. actual**: e.g. more than 10% of scheduled activities not detected by receivers, e.g. due to noncompliance of patients parallel activities within a single location

CarePath = 61 Visits: Clinical + Operational



1

ABC 3G 13:30

Elaine

Next: CT Scan At: 15:30

Geting there

Time left: about 5 hours

CT Scan
Blood test
Exam 1
Study
Blood test
Exam 2
X-Ray
Blood test
X-Ray

Idle state:
Seeing the planned schedule before anything started
The next activity highlighted up front

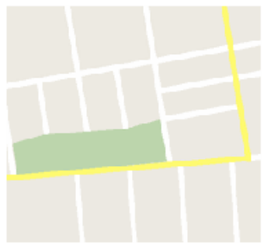
2

ABC 3G 13:30

Elaine

Next: CT Scan At: 15:30

Go to: Clinical center 10



Back to full schedule

Next activity screen

3

ABC 3G 15:16

Elaine

Next: Blood test At: 16:15

Go to: Clinical center 6

Back to full schedule

Change screen on button click

Next activity screen:
Switch to it is automatic after previous activity was completed.

4

ABC 3G 15:16

Elaine

Next: Blood test At: 16:15

Geting there

Time left: about 4 hours

CT Scan ✓
Blood test
Exam 1
Study
Blood test
Exam 2
X-Ray
Blood test
X-Ray

Idle state:
Seeing the planned schedule after one activity was completed.
The next activity highlighted up front

5

ABC 3G 15:42

Elaine

Waiting for blood test results

--:---

More info

Time left: about 4 hours

CT Scan ✓
Blood test ✓
Exam 1
Study
Blood test
Exam 2
X-Ray
Blood test
X-Ray

Idle state:
Seeing the planned schedule after two activities were completed.
The next activity highlighted up front but

6

ABC 3G 15:42

Elaine

Waiting for blood test results

--:---

More info

Time left: about 4 hours

CT Scan ✓
Blood test ✓
Exam 1
Study
Blood test
Exam 2
X-Ray
Blood test
X-Ray

Idle state:
Seeing the planned schedule after two activities were done
The next activity highlighted up front but time unknown.
Selecting a later activity to get more details

7

ABC 3G 15:42

Elaine

Later: Exam Planned at: 17:00

Location at: Clinical center 9a

Back to full schedule

Later activity screen

8

ABC 3G 16:00

Elaine

Blood test results back

Next: Exam 1 At: 16:30

Getting there

Time left: about 4 hours

CT Scan ✓
Blood test ✓
Exam 1
Study
Blood test
Exam 2
X-Ray
Blood test
X-Ray

Idle state:
Seeing the planned schedule after two activities were completed.
Status of recent activity updated.
The next activity highlighted up front with time and location.

On Data-Based Research & Teaching (Practice)

- **Operational Data** challenges: get, incomplete, corrupt; integrate w/ full carepath (e.g. from onset of symptoms, clinical)

On Data-Based Research & Teaching (Practice)

- **Operational Data** challenges: get, incomplete, corrupt; integrate w/ full carepath (e.g. from onset of symptoms, clinical)

  $\Rightarrow$ Must go beyond the prevalent paradigm (in OR/IE/OM):

  - A single researcher (with a PhD student) obtaining a small data set for a single research project, with little “clinical" interaction

    $\Rightarrow$ Unprofessional data science, no learning across generations, no sharing among researchers, irreproducible research; irrelevant

- Ultimately, **Research Labs** will become necessary (hence must be funded!): offering universal access to data, ServNets

On Data-Based Research & Teaching (Practice)

- **Operational Data** challenges: get, incomplete, corrupt; integrate w/ full carepath (e.g. from onset of symptoms, clinical)

  $\Rightarrow$ Must go beyond the prevalent paradigm (in OR/IE/OM):

  - A single researcher (with a PhD student) obtaining a small data set for a single research project, with little “clinical" interaction

    $\Rightarrow$ Unprofessional data science, no learning across generations, no sharing among researchers, irreproducible research; irrelevant

- Ultimately, **Research Labs** will become necessary (hence must be funded!): offering universal access to data, ServNets

- Data-based Research: Tradition in Physics, Chemistry, Biology; Psychology, Transportation (Science), (Behavioral) Economics

- Natural, needed and **feasible** in Operations Research & Management, Service Engineering; BPM, Process Mining

On Data-Based Research & Teaching (Practice)

- **Operational Data** challenges: get, incomplete, corrupt; integrate w/ full carepath (e.g. from onset of symptoms, clinical)

  $\Rightarrow$ Must go beyond the prevalent paradigm (in OR/IE/OM):

  - A single researcher (with a PhD student) obtaining a small data set for a single research project, with little “clinical" interaction

    $\Rightarrow$ Unprofessional data science, no learning across generations, no sharing among researchers, irreproducible research; irrelevant

- Ultimately, **Research Labs** will become necessary (hence must be funded!): offering universal access to data, ServNets

- Data-based Research: Tradition in Physics, Chemistry, Biology; Psychology, Transportation (Science), (Behavioral) Economics

- Natural, needed and **feasible** in Operations Research & Management, Service Engineering; BPM, Process Mining

A working model for Research, Teaching & Practice Support: Technion IE&M SEELab

The Technion SEE Center / Laboratory

Data-Based Service Science / Engineering

Data-Based Creation of ServNets: some Technicalities

- **ServNets** = QNets, SimNets, FNets, DNets
- **Graph Layout**: Adapted from but significantly extends Graphviz (AT&T, 90's); eg. *edge-width*, which must be restricted to *poly-lines*, since there are "no parallel Bezier (Cubic) curves ($B_n(p) = E_p F[B(n,p)], 0 \leq p \leq 1$)
- **Algorithm**: Dot Layout (but with cycles), based on Sugiyama, Tagawa, Toda ('81): "Visual Understanding of **Hierarchical** System Structures"

Data-Based Creation of ServNets: some Technicalities

- **ServNets** = QNets, SimNets, FNets, DNets
- **Graph Layout**: Adapted from but significantly extends Graphviz (AT&T, 90's); eg. *edge-width*, which must be restricted to *poly-lines*, since there are "no parallel Bezier (Cubic) curves ($B_n(p) = E_p F[B(n,p)], 0 \leq p \leq 1$)
- **Algorithm**: Dot Layout (but with cycles), based on Sugiyama, Tagawa, Toda ('81): "Visual Understanding of **Hierarchical** System Structures"
- Draws data directly from **SEELab** data-bases:
  - Relational DBs (Large! eg. USBank Full Binary = 37GB, Summary Tables = 7GB)
  - Structure: Sequence of events/states, which (due to size) partitioned (yet integrated) into days (eg. call centers) or months (eg. hospitals)
  - Differs from industry DBs (in call centers, hospitals, websites)