The Technion Center for

Service Enterprise Engineering (SEE)

ie.technion.ac.il/Labs/Serveng

Faculty of Industrial Engineering and Management
Technion
Israel

Report 2007–2010

Table of Contents

1. Introduction

The **Center for Service Enterprise Engineering** (**SEE**) was established in February 2007, within the Faculty of Industrial Engineering and Management, Technion, through the generous support of **Hal and Inge Marcus**. The goal of SEE has been the support of research and teaching in the area of Service Systems. Specifically, SEE has been contributing to the development of engineering and scientific principles that support modeling, design and management of Service Enterprises, for example financial services (banking, insurance), health services (hospitals, clinics), government and tele-services (telephone, internet). SEE's main activity has been through the SEE Laboratory (SEELab), specifically creating, designing, maintaining and analyzing an accessible repository of resources and data. The original data-sources were telephone call-centers, but the scope has been now expanded to cover also hospitals, and first steps have been taken to expand into other services such as internet, emergency and fast-food services.

In the present report, we shall discuss interchangeably the SEE Center and the SEELab, referring to both as SEE. SEE's activities span the wide spectrum from research through teaching to practice. As indicated, the central goal of SEE is to create knowledge and disseminate it worldwide. SEE's original budget was planned for four years, and January 2010 is its last month of operation under this budget – thus, in this report, we are summarizing SEE activities over its four years of existence.

2. Contents

We start with describing the staff members of SEE—research associates and students. We then proceed with our SEE philosophy—that of integrating Science, Engineering and Management, in the context of service systems. Next, we describe SEE's support of research, teaching and applications, then its international outreach, which is followed by a more technical description of SEE-developed systems—past, ongoing and soon-to-be. We conclude with a list of references to all the material that has been generated in or through SEE. (This includes a few items that had been created in SEE's predecessor, the Technion Statistics Lab, which have contributed to SEE material.)

The concluding list of references is divided into the following self-explanatory parts: SEE reports (data analysis and manuals), research documents, teaching-related activities; and SEE-related research proposals, grants, honors, conferences/lectures, graduate students and theses, undergraduate projects, research in progress, and OCR output. (OCR = Open Research Collaboration, is an IBM initiative to fund joint Industry-Academia research; it has been funding a joint research project of IBM Haifa Research Labs, the Rambam Hospital and Technion IE&M, in which SEE has played an important role; see Section 5.4.)

3. SEE Staff: Permanent and Ad Hoc

The founders of SEE are Professors **Paul Feigin** and **Avishai Mandelbaum**, both from the Faculty of Industrial Engineering and Management. Presently, Prof. Feigin is serving as the Technion's Senior Vice-President; hence Prof. Mandelbaum presently acts as the sole Academic Director. Prof. Mandelbaum also served, for two years, as the academic director of the OCR project, which has been administered through and supported by SEE.

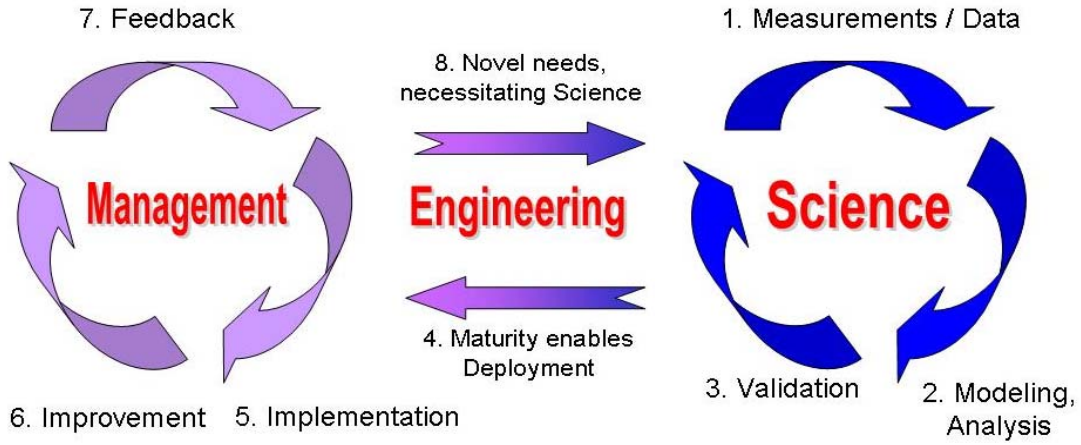
SEE employs four full-time researchers and programmers: Dr. **Valery Trofimov**, who is heading R&D, is the designer of SEE databases and applications; **Igor Gavako**, who is SEE's chief programmer; **Katya Kutssy** (MSc), who is a programmer and application-developer; and **Ella Nadjarov**, who is in charge of SEE administration and statistical-development. **The SEE Team won the 2010 Technion Excellence Award for Research.**
In addition to the permanent staff, several students, graduate and undergraduate, work regularly at the lab: **Polyna Khudyakov** (PhD student), did her thesis on SEE-related topics and supervised undergraduate projects – Polyna is now a post-doc at Harvard University; **Tirza Lauterman** (MSc

student) has been the administrative manager of OCR; and Abir Koren, Ori Plonsky and Ofri Rom (BSc students), who are members of the IE&M Honor Program. (Ofir is now continuing to an MSc degree.) On top of these "regulars", numerous graduate and undergraduate students, as well as research colleagues (locally and worldwide) have been involved in SEE activities, either regularly as part of their research or via ad hoc projects – more details will be provided in the sequel.

4. SEE Philosophy: Perpetual Multi-Disciplinary Cycle of Science, Engineering and Management

SEE data serves as an integrator of Science, Engineering and Management. In other words, SEE activities are **science**-based, with direct implications to the **engineering** and **management** of service enterprises. The following figure helps in explaining these latter assertions.

The Perpetual Cycle of Science, Engineering and Management



The classical scientific cycle (right) starts with measurements that gather data, which is then used to create models: mathematical, simulation, sometimes merely conceptual models. The models are next validated, which might lead to further measurements and model refinements, continuing within the Scientific Cycle until a maturity level is reached that enables deployment.

Service Engineering (middle) now transforms the novel Science thus created into design principles and tools, the latter typically culminating in software. Here, management (left) takes over, implementing these new designs; the designs hopefully lead to improvements, and feedback from the application often generates novel needs. These needs might necessitate further models which, in turn, calls for further measurements, leading one back into the Scientific Cycle (back to the right), and so on.

This scientific paradigm, of "Measurement, Modeling, Validation", is axiomatic in the natural sciences (Physics, Chemistry, Biology), but it is definitely a newcomer to the "Science of the Artificial", as Herbert Simon called it—that is the Science of man-made systems, such as Service Systems. The latter in fact require a paradigm that is more complex than Simon's, in that problems from service systems often call for multi-disciplinary solutions. Indeed, Industrial Engineering (IE), Operations Research (OR), Information Systems (IS), Operations Management (OM), and Human Resource Management (HRM), all must often be jointly applied to address "service-triggered" problems. But, in contrast to the Physicists, Chemists and Biologists, researchers in IE, OR, IS and OM have not been brought up with a "scientific paradigm" thinking; HRM researchers lack engineering skills; and all researchers lack experience in integrating science, engineering and management, as in the above figure. This necessitates a change in the research agenda of all

mentioned fields, as far as their applications to service systems are concerned, certainly those with significant human interactions.

The dual challenge for Service Science/Engineering, that of a multi-disciplinary integrated paradigm, raises high barriers to entry. The SEELab plays a unique key role in lowering these barriers, by enabling the various disciplines, or better yet their integrator, the Service Engineer/Scientist, to decipher, learn and master the "language of data"—this language is simply a prerequisite for solving "service problems" and doing "service research", not to mention the archiving and dissemination of the information and knowledge thus generated.

5. SEE Support of Research

Being a research laboratory within a research university naturally renders research as the primary goal of SEE. In concert with this, SEE has been providing empirical foundations and support for theoretical research, involving researchers and graduate students. It has taken some time for the research community to realize the potential of SEE resources, but the seeds have been planted and concrete promising outcomes are in the pipelines.

5.1 Theoretical Research

Specifically, SEE data has been incorporated into research that is both general in nature: e.g. testing the validity and accuracy of generic forecasting methods; or application-focused: e.g. staffing based on robust optimization methods, estimating customer impatience over the phone, or modeling the flow of calls through a call center's IVR (Interactive Voice Response, namely the phone's answering machine). As is often the case in theoretical research, a model that is developed for a particular use finds application in another: e.g. the IVR model turns out to be similar to a model of a hospital ward, and it is the data from the application areas (available in SEE) that gives credit to this similarity, thus making it useful. (The IVR model was developed in an M.Sc. thesis and the Ward's model in a Ph.D., both at the Technion.)

5.2 Empirical Research

As already described above, empirical research has been mostly lacking from the research agenda of traditional Industrial Engineering, Operations Research and Operations Management, as well as from the emerging disciplines of Service Engineering and Science. However, this state of affairs is starting to change, as manifested by conferences that are devoted to, say, Empirical Operations Management, or academic positions that require training in both theoretical and empirical research. SEE plays a central role (to be modeled after) in this ongoing change, and its significance will only increase in the future.

The Laws of Congestion: We outlined the role of SEE as an enabler and facilitator of the Scientific Paradigm. Related to this, we have started to use SEE data to validate or refute theory, specifically the Theory of Congestion relevant to call centers, hospitals and websites. For example, we are analyzing when Little's Law is valid, when it is not, and why not. SEE data has triggered the development of new theory and further empirical analysis. For example, service durations and customer patience in call centers were found to be non-exponential (exponentiality had been the traditional hypotheses), which triggered the need for a theory of many-server queues with general services and patience. Such a theory has started to develop and in fact has already come a long way. As a last example, SEE data from emergency departments has given rise to ample unexplained or only partially-explained phenomena, which is now written up.

5.3 Graduate Students

SEE has been supporting research of graduate students in many ways. SEE has also greatly enhanced these students' learning experience, by providing them with an environment and tools that enable data-penetration, in depths and breadth that are unavailable elsewhere. Some of the students that have directly enjoyed SEE resources are (Technion students unless mentioned otherwise): Noa Zychlinsky (M.Sc., ongoing), Rony Gehbali (M.Sc., ongoing), Itamar Zaied (M.Sc., ongoing), Asaf Zviran (M.Sc., ongoing), Polyna Khudyakov (Ph.D., presently at Harvard as a post-doc), Galit Yom-Tov (Ph.D., presently at Columbia University as a post-doc), Yariv Marmor (Ph.D., presently at the Mayo Clinic as a post-doc), Michael Reich (M.Sc., ongoing), Daniel Nevo (M.Sc., ongoing at the Hebrew University), Arik Senderovic (M.Sc., ongoing), Yair Goldbert (Ph.D. 2009, the Hebrew University, spent a year at SEE and used its data in his thesis; using now SEE data in his research), Shimrit Maman (M.Sc. 2009), Yulia Tseytlin (M.Sc. 2008), Zohar Feldman (M.Sc. 2008, continuing for his Ph.D.), Luba Rosenshmidt (M.Sc. 2007) and Sivan Aldor (M.Sc. 2006; Ph.D. presently at Wharton – has used SEE data and support for both graduate degrees). The SEELab has also supported research of graduate students abroad, some actually visiting SEE to learn how to use its data—this will be described in Section 8.

5.4 The OCR Project

The Rambam hospital in Haifa, IBM Haifa Research Labs and Technion's IE&M, have partnered in a multi-disciplinary project, aimed at significant improvement across the clinical, operational and financial dimensions of hospital patient-care processes. The project, which is now in its third year, has been funded within IBM's OCR framework, where OCR stands for Open Research Collaboration (the latter between Industry and Academia). The impact of this wide-arching project encompasses all partners, and is both practical and foundational. Several improvement efforts, already completed, are expected to significantly improve metrics, such as patient length-of-stay and a number of clinical errors. Furthermore, one of the major long-term goals of the project is to provide a system for comprehensive real-time monitoring and control of patient care processes, thereby potentially revolutionizing the design and management of these processes. While much work remains, significant progress towards this goal has been made. When implemented, the vision is that data from such a monitoring system will be integrated into the SEE data-repositories, as well as into a planned SEE online simulator of an Emergency Department.

This OCR project has also generated novel methods that can be applied to other service domains. Examples include: Generic model-based data completion techniques, which can be used for monitoring and process improvement scenarios where only partial data is available in existing IT systems; and the creation of models combining quality (e.g. clinical) models and operational considerations, which can be applied to any service domain with sensitivity to response time. Furthermore, this work has a significant academic impact, resulting in publications, new course materials, creation of new courses, and being the basis for several graduate theses. Finally, this work has also impacted IBM's healthcare research plans and offerings.

5.5 Online Measurements of Hospital Operational Performance: RFID and Alternatives

How will hospital data be collected, and possibly used online? This is a significant challenge in the hospital environment, where data-collection is taking low priority, and rightly so, relative to the routine—indeed, hospital routine often entails saving lives! It follows that data must be collected; equivalently, performance must be measured, automatically. Here tracking technologies, such as RFID and its competitors (Infra-red, Ultra-wide-band, Ultra-sound), come to the rescue.

Let us call OTT any Online Tracking Technology that is capable of tracking, online, the flow of patients and the location + utilization of equipment. Implementing such an OTT in a hospital will dramatically increase management capabilities and expand research potential. The hope is thus to

implement such an OTT, and direct the data it collects to the SEE. Initial attempts have been carried out in two directions.

First, in May 2010, RFID technology was implemented at the Rambam hospital in Haifa, in order to track, within the hospital, a nation-wide drill of a Mass Casualty Event. The results were only partially successful and lessons are still being learned. But even this partial success is already significant and promising. Indeed, it has already given rise to an undergraduate project of honor students, as well as to an MSc thesis, with both enjoying empirical support by SEE.

Second, a connection of SEE with 3 hospitals in Singapore has been established. These hospitals have deployed, for several years already, RFID-based tracking of patients and equipment. More specifically, in two hospitals, tracking starts upon hospitalization and stops upon a patient's release, with the exception of the Operating Theatre (OT) if relevant. At the third hospital, tracking is confined to only the OT. (Significantly, none of the hospitals tracks patients within its Emergency Department (ED), since the deployed OTT is not accurate enough for such an implementation.) The hope is that these Singapore connections materialize into the transfer of hospital data from Singapore to SEE, which will open up ample opportunities across ample dimensions.

5.6 Further Ongoing Research

There are presently several ongoing research projects, fusing theoretical and empirical research, in which SEE and its data play a central, sometimes leading, role. We now present a partial list of these projects – in all, Prof. Mandelbaum is the SEE research connection, while SEE researchers provide the empirical support required.

- Gans, N., Shen, H. *"Heterogeneous Servers: Theory- Based Empirical Analysis"*. NSF-supported research, in which SEE data is used for analyzing operational performance heterogeneity of telephone agents. Such heterogeneity is manifested through agents' differing learning curves, and is amplified by call centers' high turnovers. The research aims at understanding the effects of these phenomena on workforce management, and then developing insights and tools that accommodate the challenges thus presented.

- Maman, S., Whitt, W., Zeltyn, S. *"QED Q's with Random Parameters"*. This work, which is based on the M.Sc. thesis of Shimrit Maman, was triggered by the need to understand and quantify the magnitude of random variability in the demand for services (in both call centers and hospitals). Such an understanding is important, since more variability goes hand-in-hand with increased staffing levels that are required to cope with it. SEE data has revealed fascinating phenomena, some of which have been incorporated into a research paper presently under preparation, and others have given rise to ongoing subsequent research.

- Gurvich, I., Liberman, P. *"Empirical Analysis of Skills-Based Routing in Call Centers: A Queueing Science Perspective"*. Skills-Based Routing (SBR) is the protocol of matching customers to telephone agents, taking into account agents' skills and customers' attributes (e.g. needs, priorities). Being challenging theoretically and important practically, the design of SBR protocols has enjoyed significant attention by the research community. All this research, however, has been lacking empirical foundations, a gap that the present research aims at closing.

- Armony, A., Marmor, Y., Tseytlin, Y., Yom-Tov, G.*"From the Emergency Department to Hospitalization and Beyond: Using Simulation, Jointly with Empirical and Theoretical Models, for the Operational Analysis of ED, IW, and Their Interface"*. A team of graduate students, all working in the area of healthcare, teamed up for an empirical analysis of hospital data, covering the care process from the emergency department to release. Academic guidance has been provided by the SEE staff, as well as by Prof. Mor Armony, of the Stern Business School at NYU. The project is still ongoing even though the students have all graduated: one works at IBM

Research Lab, one at Rafael, and the two others are postdocs in the U.S. (Columbia University and The Mayo Clinic). Needless to say, the students' SEE experience has truly helped them to secure the attractive positions they are now enjoying.

- van Leeuwaarden, J.S.H., Borst, S., Boxma, O., Zwart, B. "*Pre-limit Results for Queues in Heavy Traffic*". Research proposal to the NWO Free Competition, Physical Sciences, The Netherlands. This Dutch proposal seeks support for a PhD student, to be directly advised by the first proposer, and consulted by the rest of the team (all from the Eindhoven University of Technology). Indeed, part of this student's training and research will be carried out at the Technion, working on SEE data, and using it to support or refute the theory to be developed. The theory is that of asymptotic analysis of many-server queues, for which call centers have provided the main trigger and continuous inspiration.

- Armony, M., Bassamboo, A., Hurvich, C., Rouba, I. "*Forecasting Waiting Times in Telephone Queues*". Many call centers provide estimations for waiting times to their customers. The goal here is to develop the theory of some simple, yet scientifically-grounded estimation methods, then test their accuracy with SEE data.

- Alon, G., Bassamboo, A., Khudyakov, P. "*Analysis of Information Provided to Customers while Waiting for Service*". SEE repositories include a unique collection of VRU (Voice Response Unit), often also referred to as IVR (Interactive Voice Response) data. Some of these VRU's are sensitive to the state of congestion, in which case they encourage customers to self-serve if congestion levels are high. The first two researchers, both of the Kellogg Business School in Northwestern University, have developed theoretical results on information to customers. Again, SEE data will be used to either support or refute this theory.

- Goldberg, Y., Reich, M., Ritov, Y. "*Improving the Forecast of Arrivals and Offered-Load with Same-Day Information*". Statistical analysis of demand for service, in the form of arrivals or offered-load, exhibits intra-day correlations that, in fact, can be exploited for improving forecasting. For example, knowing the arrivals from 8am to 10am on a certain day can improve a previous-day's forecast of that day's afternoon arrivals. As part of Goldberg's Ph.D. thesis, under the guidance of Prof. Ritov at the Hebrew University, Spline-based methods were developed for improvements, as described above. These methods have been successfully tested on SEE data. Some of the work is also relevant to the M.Sc. thesis of M. Reich, jointly supervised by Professors Ritov and Mandelbaum.

- Aksin, Z., Ata, B., Chen-Lin, S., Emadi, S. *"Structural Estimating of Callers' Delay Sensitivities in Call Centers Under Delay Announcements"*. Theoretical and empirical research on delay announcements for customers awaiting phone services.

- Xiaobo, Z., Qi-Ming, H., Hao, Z., Nitzan, Y. *"Human Aspects in Queues"*. Nitzan Yuviler, a Technion graduate, delivered a workshop at Tsinghua University, Beijing, China. In that workshop, Nitzan taught the participants (faculty and gradate students at the IE Department) the basics of SEE data, including SEEStat. This workshop, with the blessing of Prof. Gavriel Salvendy (IE Head), established the foundation of a joint research venture, involving the IE Deputy Head (Xiaobo), his colleague from Canada and a Tsinghua graduate student.

- National University of Singapore (Decision Sciences at the Business School), jointly with the SEELab, are exploring options for cooperation that will help create a Singaporian SEELab, focusing on Health-Care. As described above, this new SEELab will draw data from unique RFID-based data repositories, available at several Singapore hospitals, which will also be shared with the Technion SEELab.

6. SEE Support of Teaching

Second to research, teaching support is SEE's next primary goal. This support has taken several forms.

6.1 The Technion Service Engineering Course

A unique Service Engineering course has been developed and taught at the Technion over the last decade. SEE's existence is taking this course to new levels. Indeed, today's students of Service Engineering can either install SEEStat on their own computers (this is SEE's interface machine), or access SEEStat via the SEEServer. In either case, students are accessing and using SEE data as part of their course homework.

The INFORMS Journal on Transaction on Education recently published (September 2010) an invited review paper of the Technion's Service Engineering course.
See www.informs.org/Pubs/ITE/Archive/Current-Issue.
The paper's title is "Service Engineering: Data-Based Course Development and Teaching," which is directly referring to the prominent role that SEE data and the SEELab, have been playing in the course—past and present. Specifically, the paper describes how SEE material has been regularly used in course material (lectures, recitations, exams); how through this data, phenomena are being explained, demonstrated and tested, and theory either supported or refuted. Examples include empirical discussions of Little's Law, Service Durations, as well as Customer Impatience over the phone. A recent example, not included in the paper, is a queueing-focused modeling of the flow of patients through a hospital's Emergency Department, which constituted part of the MSc thesis of the course's Teaching Assistant (course generation 2011W).

Significantly, the Service Engineering course has been attended by Technion research partners. For two such examples: several members of the IBM Haifa Research Lab have taken the course, some going beyond merely listening to lectures but rather performing all the required homework (which is a significant undertaking); and the Operations Manager of Rambam's Imaging Unit attended the course as well. These students then maintain close relations with SEE activities beyond the course. For example, one of IBM's members is now pursuing a graduate degree at the Technion; and the Rambam manager is consulting a graduate student in his thesis and advising a colleague on SEE-related research.
The Service Engineering course website is accessible to researchers worldwide, and it is frequently visited and its material used; see http://ie.technion.ac.il/serveng especially Recitation 6, and homework 6 in http://iew3.technion.ac.il/serveng/Homeworks/HW5_2010S.pdf (though SEE material is used in many of the classes as well).

6.2 Mini-courses on Service Engineering

A mini-course of Service-Engineering has been taught, in recent years, at Stanford, Columbia and Wharton. The course was attended by PhD students and faculty. Thus, SEE-generated knowledge, as well as SEE resources, is being spread worldwide.

Seminars on SEE-related topics: There have been two semester-long SEE-related seminars at the Technion, one dedicated to Call Centers and the other to Healthcare Call-Centers. These seminars were attended by students, faculty and outside visitors (from industry).

Data for Other Courses: SEE has provided data for other data-hungry courses at the Technion: Data-Mining and Time-Series. Both courses require fresh data for homework assignments and students' projects, and SEE data and its staff are ideally suited to provide such support.

6.3 Undergraduate Projects

During the 3rd or 4th year of studies, every undergraduate IE student must take a project-course, supervised jointly by industry and academia. In recent years, the SEELab has played an important role in many such projects. Specifically, these projects have been performed jointly with SEE partners (banks and hospitals), who provide the SEELab with its data. Examples of such projects include:

- Analysis of call-duration and call-structure at a bank's call center.
- Fairness vs. Efficiency in the routing of patients from the emergency department to internal wards.
- Fusing operational data with financial data towards profit prediction at a call center.
- Operational models and data infrastructure, in support of management and control of a Mass-Casualty-Event at a hospital.

In all the above examples, and others, SEE members have actively participated in project support and supervision, thus providing expert advice and a data-bridge between the students and their partner companies. There have been additional projects that took place in companies or institutions that have no data-relations with SEE. However, their data was imported to SEE and interfaced with SEEStat, so that the students and their industry partners could view and analyze the data in a mutually beneficial manner. One such instance was a project with the Israeli Police Call Center (100)—consequently, a process was initiated by which Police data will be transferred to the SEELab, much in the same way as banks and hospitals.

7. SEE Support of Practice (Service Engineering)

SEE's industry partners have been enjoying and benefitting from their collaboration with SEE. For a start, the import of data to SEE necessitates a thorough understanding of the data-structure, which included data cleaning, discoveries of inconsistencies, etc. Thus, at the end of the import process, the data-sources themselves have typically improved significantly.
SEE members are also consulted, by SEE partners, on problems that can find their solutions in SEE data. As an additional example, groups from a bank- and hospital-partner of SEE were hosted at the Technion for a full afternoon. The meetings centered on lectures by SEE members, on interesting findings from these partners' SEE data, which had been imported to the SEELab. These meetings were declared a major success and, since then, have helped all parties involved in their activities.

It is important, in this context, to elaborate on the present model of SEE-Industry cooperation, which takes intense commitment and efforts from both sides. Indeed, the Industrial partner is committed to data-collection, data transfer and continuous maintenance of the process, which is no small endeavor. The SEELab, on its side, is accepting the data, checking it for inconsistencies, cleaning it, and absorbing it so that it can be interfaced with lab resources, especially SEEStat.

The resources required from SEE to absorb industry data are so significant that only one or two partners at the most can be accommodated simultaneously. (This is a constraint that is important to account for when considering, for example, commercialization of SEE resources.) Indeed, presently, SEE has two active data-partners: a mid-size commercial Israeli bank and a large hospital. Consider the SEE-bank relation, as a concrete example. A mutual decision to transfer data from the bank to SEE was made at the end of 2007. Following a tedious process that involved the resolution of ample legal and technical obstacles, the first "bit" of data was transferred to SEE in mid-2009, about a year and a half later. Stabilizing the data transfer process took one additional year, and presently SEE has about 1 year's worth of data from the call center of that bank.

8. SEE (International) Outreach

SEE's reputation is building up, and its resources are increasingly acknowledged and utilized. Examples were already given in Section 5.6, within SEE's ongoing research. A sample of additional outreach activities is the following:

- Lecturing about SEE and its databases, including leaving SEE Data (via a DVD or hard disk): USA (Columbia, Stanford, UPenn, NYU); Turkey (Koc, Bogazici); Italy (LUISS), Germany (Berlin, Stuttgart, Ulm), Holland (Amsterdam, Eindhoven), China (Beijing, Shanghai), Argentina (Buenos Aires).
- Hosting a PhD student (Yair Goldbert) from the Hebrew University during 2008–2009; in the final year of his studies, the student learned to use SEE databases, which have been incorporated into his thesis. He is now a postdoc at the University of North Carolina, planning to cooperate with Prof. H. Shen on SEE-based research.
- Hosting a PhD student (Junfei Huang) from the National University of Singapore (NUS), between 10/2010 and 1/2011. The student is learning to work with SEE data, with the goal of combining it with theoretical research. Furthermore, the expertise acquired will be exported to NUS, for the benefit of students as well as faculty. Additional visits of PhD students are planned for Dutch and US graduate students (jointly with their advisers).
- An excellent Technion graduate (Nitzan Yuviler) is spending the year of 2010 in China with her husband, who is enjoying a Chinese fellowship. Prior to her trip, Nitzan had spent time at the SEELab, learning how to use its programs and access its data. The idea was to try and export this knowledge to China, and it has been working wonderfully. Indeed, Nitzan was hosted by the IE department at Tsinghua University (China's MIT), and conducted a workshop for faculty and graduate students. To quote Prof. Gavriel Salvendy, who is heading this IE department: "Nitzan was outstanding and has made the Technion proud. The data is outstanding and there are endless opportunities for seminal high impact contributions here. I sincerely hope that under Professor Xiaobo's leadership at our end and yours at the Technion, we will jointly move to the next level." Nitzan's visit has given rise to two opportunities: first, a joint research program is now in planning, between Tsinghua IE and the SEELab, with a possible submission of a proposal to the Chinese NSF; and a visit of Nitzan to Hong-Kong is now contemplated as well.
- Hosting researchers from Israel and abroad for the purpose of demonstrating SEE capabilities. For example, Prof. H. Takagi, who is managing an academic program in Service Science and Engineering at Tsukuba University, came for a 2-day visit from Japan, exclusively to learn about our SEE data-resources and research activities. There have been many additional learning-visits from researchers and practitioners, from both USA and Europe.
- Hosting foreign exchange students (undergraduates), for example a student from Germany during the summer of 2009.
- Cooperating in and contributing to research, ongoing and proposed, in the U.S. (see Section 5.5).
- Providing a home for OCR data (OCR = Open Collaborative Research, jointly The Rambam Hospital, IBM Research Labs in Haifa and IE&M at the Technion). Initial contacts have been initiated in Philadelphia, New York and Singapore, exploring the possibility of hosting data from U.S. and Singapore hospitals at SEE.
- In May 2010, the Technion hosted an important international conference: MSOM 2010 and its Special Interest Groups – SIG (http://msom.technion.ac.il/index.php?num=1). The SEE Center was a formal sponsor of this conference, and it played an important role in its program (and ultimate great success):

During the SIG on Services, Prof. Mandelbaum delivered a lecture on "Empirical Adventures in Call Centers and Hospitals", based on SEE data. The lecture was followed by a workshop for 15

international participants, who went over a self-teaching tutorial, prepared by the SEELab, that taught them, online, to work with SEEStat on SEE data. This workshop was a great success, and it inspired many additional MSOM participants to practice this tutorial at the SEELab, on an individual basis. The tutorial can be accessed at
http://ie.technion.ac.il/serveng/References/MSOM_2010/SEEStat%20Workshop%20Tutorial%20_solution.pdf

Two full sessions, which included 7 lectures, were devoted to research that is based on SEE data. (It is noteworthy that MSOM is a competitive conference, namely these lectures had been pre-submitted to the conference and chosen among many others.) The lectures can be accessed at
http://ie.technion.ac.il/serveng/References/MSOM_2010/MSOM_Main_Page.htm.
Their titles and authors are as follows:

- *Emergency Department (ED): The Case for Service Engineering*. Marmor Yariv.
- *On the Accuracy of Delay-History-Based Predictors in Large Call Centers*. Armony Mor.
- *Skill Based Routing: Data-Based Review and Research Prospects (Part II: Staffing and Routing)*. Gurevich Itay.
- *Uncertainty in the Demand for Service: The Case of Emergency Department and Call Centers*. Zeltyn Sergey.
- *Queueing Systems with Heterogeneous Servers: On Fair Routing from Emergency Departments to Internal Wards*. Tseytlin Yulia.
- *Design and Inference of a Call Center with an Answering Machine (IVR)*. Khudyakov Polyna.
- *The Erlang-R Queue: Time-Varying QED Queues with Reentrant Customers in Support of Healthcare Staffing*. Yom-Tov Galit.

SEE data has been used, with acknowledgements, in research publications worldwide. Examples include testing methods for arrival forecasting, phase-type fitting, (im)patience estimation, and more. We now list these publications. More details including Abstracts and links, will appear at the end of this document.

- R. Atar. *A Diffusion Regime with Non-degenerate Slowdown*, June 2010. http://webee.technion.ac.il/people/atar/NDS-rev.pdf
- K. Kearton. *Correlating Temporal Rules to Time-Series Data with Rule-Based Intuition*, March 2010. http://edocs.nps.edu/npspubs/scholarly/theses/2010/Mar/10Mar_Kearton.pdf
- A. M. De Livera, R. J. Hyndman. *Forecasting Time Series with Complex Seasonal Patterns using Exponential Smoothing*, December 2009. http://robjhyndman.com/papers/ComplexSeasonality.pdf
- D. Bestimas, X. Vinh Doan, *Robust and Data-Driven Approaches to Call Centers*, November 2008. http://web.mit.edu/dbertsim/www/papers/Robust%20Optimization/DataDriven%20and%20Robust%20Optimization%20Approaches%20to%20Call%20Centers.pdf
- H. Shen, J. Z. Huang. *Interday Forecasting and Intraday Updating of Call Center Arrival,* Manufacturing & Service Operations Management, Vol. 10, No. 3, summer 2008, pp. 391–410. http://msom.journal.informs.org/cgi/content/abstract/10/3/391
- J. W. Taylor. *A Comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center*, Management Science, February 2008, Vol. 54, pp. 253–265. http://users.ox.ac.uk/~mast0315/CallCenterFcstComparison.pdf
- M. Concepcion Ausin, Michael P. Wiper, Rosa E. Lillo. *Bayesian Prediction of the Transient Behaviour and Busy Period in Short and Long-tailed GI/G/1 Queueing Systems*, January 2008. http://www.udc.es/dep/mate/Dpto_Matematicas/Investigacion/ie_publicacion/Ausin.Wiper.Lillo.pdf
- Anders Gorst-Rasmussen and Martin B. Hansen. *Asymptotic Inference for Waiting Times and Patiences in Queues with Abandonment*, July 2007. http://www.math.aau.dk/fileadmin/user_upload/www.math.aau.dk/Forskning/Rapportserien/R-2007-14.pdf
- A. Thümmler, P. Buchholz, M. Telek. *A Novel Approach for Phase-Type Fitting with the*

EM Algorithm, July 2006. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1673383
- J. Weinberg, L. D. Brown, J. R. Stroud. *Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data*, June 2006.
- J. Z. Huang, N. Liu, M. Pourahmadi, L. Liu. *Covariance Matrix Selection and Estimation via Penalized Normal Likelihood*, Biometrika (March 2006), Vol. 93, No. 1, pp. 85–98. http://biomet.oxfordjournals.org/content/93/1/85.full.pdf
- H. Shen, L. D. Brown. *Nonparametric Modeling of Time-varying Customer Service Times at a Bank Call Center*, January 2006. https://mercury.smu.edu.sg/rsrchpubupload/7292/c2006-1-1.pdf
- Sandjai Bhulai, Wing Hong Kan, Elena Marchiori. *Nearest Neighbour Algorithms for Forecasting Call Arrivals in Call Centers*, August 2005. http://www.cs.ru.nl/~elenam/WS2005-12.pdf
- H. Shen, J. Z. Huang. *Analysis of Call Centre Arrival Data using Singular Value Decomposition*, Applied Stochastic Models in Business and Industry, May 2005. http://onlinelibrary.wiley.com/doi/10.1002/asmb.598/pdf
- E. Chassioti, D. J. Worthington. *A New Model for Call Centre Queue Management*, The Journal of the Operational Research Society, Vol. 55, No. 12 (Dec. 2004), pp. 1352–1357. http://www.jstor.org/stable/pdfplus/4101854.pdf?acceptTC=true
- Yen Chu Cheng. *Analysis of Call Center Data*, April 2004. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1002&context=wharton_research_scholars
- L. D. Broun, L. H. Zhao. *A Test for the Poisson Distribution*, Sankhya: The Indian Journal of Statistics, 2002, Vol. 64, Series A, Pt. 3, pp. 611–625. http://sankhya.isical.ac.in/search/64a3/64a3037.pdf
- Søren Asmussen, Olle Nerman, Marita Olsson. *Fitting Phase-type Distributions via the EM Algorithm*, Scandinavian Journal of Statistics, Vol. 23, No. 4 (Dec. 1996), pp. 419–441. (This research uses data that was later absorbed into SEE repositories.) http://www.jstor.org/stable/4616418

9. SEE Data and Software

SEE has been developing an environment for accommodating data from service enterprises. The environment consists of software modules, which are convenient to group into the following system components:

1. Data repository.
2. Core software components.
3. Data processing tools.
4. Analysis and modeling tools.
5. User interface.

We now describe each of these components in some detail.

9.1 Data Repository

Service center data is stored mainly in relational databases or in other formats, for example, in click stream files for Internet sites. The structure of these databases varies among service centers, depending mainly on the commercial software used by the center. Our precondition for creating a data repository of service was the design of the *uniform* representation of data, accommodating a wide variety of service centers.

In SEE, we first developed a data-model for call center analysis, entitled **DataMOCCA**, which is as independent as possible of the particular format of its source data. (DataMOCCA stands for Data MOdel for Call Center Analysis**.)** This data-model, initially designed for call center data, has been

extended to accommodate a generic service center, and the concept has been successfully used to integrate data from hospitals and internet sites.

The main challenge for the creation of such a data repository is the huge volume of data. For example, the largest database in our data repository has the size of around 50 GB. Even the most powerful SQL servers are too slow to extract data from a database of such a size, not to mention if a time-series for a long period of time is needed within a few seconds.

The objective was to create a system that could work on a PC, which is to be used by researchers and students outside SEE, and provide real-time access to data, with maximal delays of a few seconds. To achieve this objective, the concept of a *database partition* was developed and implemented.

Call center databases are organized as a system of daily databases. This allows the splitting of huge files into relatively small databases, of up to around 100 MB each. The information for each call is stored in the database according to the day in which it was initiated. The duration of the call is less than one day, and there are typically only a few calls that start before and end after midnight. These have been processed for the next day by creating small tables in the database for the next day. Click stream database is structured in the same way. However, for health care databases, a patient's hospitalization could last weeks and even months. Information for the entire visit is located in the database for the month in which the visit was initiated, but it is also needed for all days until the visit ends. To solve this problem, the health care databases have been structured on a monthly basis. If the visit started before a given month, additional tables are added to subsequent monthly databases, providing information relevant to that month. This modification affected only the process of extracting daily data from monthly databases. The successful integration of health care data into DataMOCCA demonstrates that this system design is general and flexible enough. It can accommodate various types of Service data and, thus, could become a standard tool for Service Science and Engineering research and application.

SEE data repository consists of *studies*. Despite the splitting of the database into dozens and hundreds of daily or monthly tables, the system works as an integrated database, using structures created at the study level, such as study calendar, dictionaries and database table descriptions. SEE now hosts three large call center studies: a large call center of a U.S. bank, which provided us with 2.5 years' worth of transaction-level data (around 218 million calls), an Israeli Telecom Company with over 3 years' worth of data (175 million calls), and an Israeli Bank with over 1 year of data (26 million calls). An additional small call center study was integrated from the Anonymous Bank research project – this was a pioneering research effort, during 1999–2003, which culminated in a joint Wharton-Technion empirical research that set the stage for the data-based research of call centers.

The SEE data repository now hosts also two health care studies: Emergency Rooms (EDs) of 5 Israeli hospitals (each with about 2 years' data); and a comprehensive database (ED, Wards, Operating Rooms, Labs, and more), still active, from a partner Israeli Hospital, spanning a period of 4 years.

As described above, in 2009, a new study was initiated at the call-center of an Israeli mid-size bank. In this study, the data-transfer process is unique in that operational data is fed into the SEELab on a **daily basis**. In other words, every night, after the bank archives its call-center daily data, that data is deposited in a "safe" at the SEELab, and then automatically incorporated (after cleaning and validation) into the corresponding SEE study. The plan, in the not-too-distant future, is to expand the flow of operational data to also cover financial data—needless to say, this planned combination of operational and financial data, which has never yet been achieved, promises endless opportunities for research (empirical and theoretical) and applications. It will also pave the way for analogous data-gathering efforts in hospitals, which will combine operational and clinical data.
SEE also contains a study from the IE&M Faculty website, based on Internet click stream data over 1.5 years. Initially, the SEE data repository was created using Access. In 2009, we switched to a self-developed binary format—this version is to be implemented during 2010. This transition to binary format has improved the time performance of our data processing by 5 to 10 fold. Additionally, the

use of custom designs of our database has opened up the possibility for developing new data processing algorithms, which could not have been implemented with Access databases.

9.2 Core Software Components

9.2.1 Class Libraries

Class libraries are essential components for the development of all other components of the SEE system. Main applications, containing visual interface and connection to Microsoft Office, must be written in one of the .NET languages (C# or Visual Basic) – we selected Visual Basic. On the other hand, computing algorithms have been written in C++. A new extension of C++, called CLI/C++, was recently developed by Microsoft—this extension provides a transparent interface between .NET and C++. In 2008–2009, the core class libraries, previously written in Visual Basic, were replaced by new libraries in CLI/C++. Fundamental classes, such as vectors, lists, matrices, etc. from these libraries, are used in C++ and Visual Basic. Other libraries are built on these core libraries, containing a variety of vector and matrix algorithms, optimization algorithms, basic statistical functions and more—these have been written in C++ and Visual Basic.

9.2.2 Interpreter

An interpreter language was developed and implemented in the user interface. It is the base of all data processing operations. It contains 279 functions, vector and matrix subscripts, control blocks (such as, if-else, for, while, etc.). A code editor with syntax checks is also available for supporting the interpreter.

9.2.3 Calculator

Results of code processing, provided by the interpreter, are executed by the calculator. This calculator is built by the code generation program, which parses the code of class libraries and creates C++ files for the calculator library (around *600 pages of source code*). The calculator is also the basis for all the data processing operations.

9.3 Data Processing Tools

With transition to the binary database, all data processing tools have been completely rewritten. A dialect of the SQL adapted to binary format, specific needs of service engineering data and partitioned database structure have been developed and integrated into software.

9.3.1 Database Procedures

Database procedures (dbprocedures) are used to build a database extension by creating new database tables from existing databases. The interface to design dbprocedures allows users to define standard SQL queries, as well as the creation of new fields, based on full functionality of the interpreter. New special dbprocedure components, named macros, have been recently introduced. They provide the possibility to build tables, which cannot be created with SQL, for example, queue status (slips and skips). The dbprocedures could be nested (like a tree), allowing the user to design complex data processing algorithms. This type of dbprocedure was used to create the customer database from daily databases of the Telecom study, containing summary information for each customer for more than one million records. Dbprocedures are stored in the study files and could be used repeatedly, over many times.

9.3.2 Data Import

As mentioned above, data sources for the SEE data repository are found in various formats. They should be converted to the DataMOCCA format. For existing studies, this was done by special conversion programs, customized for each study. A significant amount of code in these studies was written for the validation of data and some corrections of it (lost or inconsistent call IDs, time

discrepancies, and the likes). An import-data tool was developed to import external files through our user interface. This tool provides the conversion and transformation of data, as well as its integration into the SEE data repository. Combined with dbprocedures, this should facilitate the process of extending the SEE data repository. Furthermore, for relatively accurate data, such imports can be performed automatically without writing additional code.

9.3.3 Summary Tables

Extracting data from databases that contain hundreds of millions of records, and then statistically processing the extracted records, would take a lot of time even for the most powerful statistical software (in the order of hours, often many hours). This is prohibitively long, rendering impossible our goal of online data analysis. SEE studies are thus designed in two levels, containing as level 2 a second-resolution data structure, namely precompiled summary tables—these are created *once* and are detailed up to 1-second resolutions (e.g. arrivals, service times, waiting times), yet they are efficient enough to support real-time few-seconds processing. Moreover, tools have also been developed to support the design and creation of new summary tables. This gives rise to an environment that is convenient for real-time data analysis.

9.4 Analysis and Modeling Tools

9.4.1 Statistical Models

In 2007, we started to develop the SEE statistical-analysis engine, which is still ongoing. We call it SEEStat. Several statistical modules are already available through SEEStat, specifically:

Distribution Fitting

The user can select several distributions from a list of distributions to fit empirical data—such as waiting-times or service-times data. In 2010, the distribution-fitting module was extended from 17 to 50 distributions—the latter collection covers essentially all distributions in practice. When fitting a distribution, the large number of candidate distributions (50) necessitates a fitting algorithm that will help the user select automatically the best fitting distributions. Such an algorithm was developed at the SEELab and integrated into SEEStat.

Fitting Mixture of Distribution

An original algorithm was developed to fit mixtures of up to seven components with various, possibly heterogeneous, distributions. Existing mixture algorithms and software packages, all require that the user provide initial values of parameters for each component (proportion, mean, standard deviation) and only one and the same distribution type (for example, lognormal) may be used for each component. The algorithm defines initial values of parameters automatically using one of several approaches: Automatic Separation, Visual Separation, Proportions Grid, and Genetic Algorithm Optimization. This application is important since many distributions in DataMOCCA studies are heterogeneous mixtures. For example, the distribution of service times in call centers is typically a mixture of at least two or three quite different components (e.g. a *main* distribution and (a) *secondary* distribution(s) of calls with short duration).

Univariate ANOVA

Capabilities for the analysis of categorical variables have been expanded. This was done because in health care studies, as well as internet click-stream studies, the majority of (explanatory) variables are categorical (Department, Internet page, Browser, etc.). One way Analysis of Variance (ANOVA) has been integrated into the analysis of categorical variables.

Survival Analysis

The following models are available:

- *Survival Curve Estimates*
  Estimation Method: Kaplan-Meier or Nelson-Aalen.
  The user can compare the empirical data from different services or different week days, month, etc.
- *Parametric Survival Curve Fitting*
  The user can select several distributions from a list of 8 distributions to fit survival data.

Smoothing Algorithms

In 2009, the statistical engine was significantly enhanced by integrating smoothing algorithms into it. The following algorithms have been integrated:

- *Kernel Density Estimation*
- *Hazard Estimation with Flexible Tails*
- *Cubic Smoothing Spline*
- *Polynomial Smoothing Spline*
- *Friedman's SuperSmoother*
- *Local Polynomial Regression Fitting*
- *Hazard Estimation using Kernel-based Methods*
- *Kernel Regression*

9.4.2 Graphs and Flow charts

Service Engineering objects (phone calls, patient sojourns in hospitals, Internet site visits) consist of several steps. For example, a customer enters the call center through the VRU; then the call is transferred to the queue, to the agent; it could go back to the VRU, or to another queue or agent, etc. Similarly, a patient in the hospital could enter the emergency department or one of the hospital departments directly, and could be later transferred to another department (e.g. ICU). And the visitor of an Internet site is also navigating through the site, visiting various pages and documents. A *graph* is a natural model for describing such movements, then for subsequent analysis of the model and the service process.

In SEE, we are now developing tools to extract data needed for the construction of graphs from databases, in order to analyze the graph using graph algorithms and to draw corresponding flow charts. Our graph library is under construction. It includes a variety of graph algorithms, as well as graph layout algorithms that are needed for graph drawing. Graph classes that are based on our core libraries have been designed as well.

9.4.3 Data-Driven Simulation

The complex reality of service centers often renders their analysis mathematically intractable—hence one resorts to *simulation*. Indeed, simulation modeling is the most viable tool for analyzing and predicting service center performance, thus providing support for management decisions at all levels. There are several problems that arise in modeling service operations. The main sources for these problems are limited data sources, unavailability of statistical models to evaluate and predict model parameters and the scope and validity of the models themselves: there is a gap between available tools for modeling and simulation on the one side and the reality of service centers on the other.

The objective is to design and create a comprehensive *system* of data sources, statistical algorithms and simulation models. The development of data-driven simulation tools is still in the exploratory phase. The prototype of the simulation software is presently undergoing testing. In fact, several modifications and upgrades of SEE software, including the transition to binary format, have been partially motivated by the needs of simulation tools.

9.5 User Interface: SEEStat

A new user interface, called SEEStat, was developed and implemented, thus replacing the previous SEE application of StatCCA (developed in 2008). SEEStat provided access to all the statistical algorithms described above, as well as to all the data processing tools. For all the new algorithms, a graphical interface was created to produce charts and statistical tables. These new algorithms, especially distribution fitting, have been proved to be extremely useful for modeling components of Service Engineering systems (queues, agent service times, etc.).

SEEStat is now in its third generation. This one is based on the binary database and extends the use of statistical algorithms to all types of data sources: summary tables, main database tables, dbprocedures results and user-imported tables. The new version of SEEStat is to be released at the end of 2010.
Finally, an online version of SEEStat has been implemented, and is now available on the SEE Internet site. There are presently (as of November 2010) 83 researchers from around the world, who are registered to the SEECenter server, in order to use SEEStat online.

References

SEE Reports

Empirical Analysis of a Call Center

1. *Empirical Analysis of a Call Center in an Israeli Commercial Company*, Khudyakov P., Mandelbaum A., Trofimov V., Maman S., Nadjharov E., Gavako I., Liberman P., Kutsi K.
2. *Empirical Analysis of a Call Center*, Mandelbaum A., Sakov A., Zeltyn S., Kutsi K.
3. *The Call Center of "US Bank"*, Donin O., Feigin P., Mandelbaum A., Zeltyn S., Trofimov V., Ishay E., Khudyakov P., Nadjharov E.
   http://ie.technion.ac.il/Labs/Serveng/files/The_Call_Center_of_US_Bank.pdf
4. *Skills-Based-Routing in US Bank*, Liberman P., Trofimov V., Mandelbaum A.
   http://ie.technion.ac.il/Labs/Serveng/files/Skills-Based-Routing_USBank.pdf
5. *The Call Center of "IL Telecom"*, Donin O., Feigin P., Mandelbaum A., Zeltyn S., Trofimov V., Ishay E., Khudyakov P., Nadjharov E.

Data processing, Data summary, User Interface

6. *DataMOCCA*, Mandelbaum A.
   http://ie.technion.ac.il/Labs/Serveng/files/DataMocca.pdf
7. *Model Description and Introduction to User Interface*, Trofimov V., Feigin P., Mandelbaum A., Ishay E., Nadjharov E.
   http://ie.technion.ac.il/Labs/Serveng/files/Model_Description_and_Introduction_to_User Interface.pdf
8. *SEEStat Beginners Tutorial*, Liberman P.
   http://ie.technion.ac.il/Labs/Serveng/files/SEEStat_Advanced_1_Tutorial.pdf
9. *SEEStat Advanced 1 Tutorial,* Nadjharov E.
   http://ie.technion.ac.il/Labs/Serveng/files/SEEStat_Advanced_1_Tutorial.pdf
10. *SEEStat Analysis - Example of Homework*, Maman S., Liberman P.
    http://ie.technion.ac.il/Labs/Serveng/files/SEEStat_Analysis Example_of_Homework.pdf
11. *Summary Tables Variable Definitions – US Bank*, Reich M., Trofimov V., Feigin P.
    http://ie.technion.ac.il/Labs/Serveng/files/Summary_Tables_Variable_Definitions-USBank.pdf
12. *Summary Tables Variable Definitions – ILTelecom*, Reich M., Trofimov V., Feigin P.
13. *STATCCA Guide II - Advanced User,* Reich M., Trofimov V., Feigin P.
14. *STATCCA Guide III - Data Extraction Facility*, Trofimov V., Feigin P., Ishay E.
15. *Design Summary Table Example*, Aldor S., Feigin P.
16. *SEEStat Installation Manual*
    http://ie.technion.ac.il/Labs/Serveng/files/ReadMe.htm

Quality Improvement and Marketing Plan

17. *Quality Improvement Plan* (in Hebrew). Sharon K., Levi N., Liberman P. http://ie.technion.ac.il/Labs/Serveng/files/Quality_Improvement_Plan.pdf
18. *SEEStat Marketing Plan* (in Hebrew), Shitrit E., Dahan D., Gutman N., Liberman P. http://ie.technion.ac.il/Labs/Serveng/files/SEEStat_Marketing_Plan.pdf

Research

19. Mandelbaum A., Momcilovic P., Tseytlin Y. *On Fair Routing From Emergency Department to Hospital Wards: QED Queues with Heterogeneous Servers*. Submitted to Management Science, 2010. http://iew3.technion.ac.il/serveng/References/fr6.pdf
20. Mandelbaum A., Feldman Z. *Using Simulation-Based Stochastic Approximation to Optimize Staffing of Systems with Skills-Based Routing*. Proceedings of the 2010 Winter Simulation Conference, December 2010. http://iew3.technion.ac.il/serveng/References/wscSAStaffing_v3.pdf
21. Gans N., Liu N., Mandelabaum A., Shen H., Ye H. *Service Times in Call Centers: Agent Heterogeneity and Learning with some Operational Consequences*. To appear in a book on the occasion of Larry Brown's 70th birthday, December 2010. http://iew3.technion.ac.il/serveng/References/agent-heterogeneity-Brown-Book-FINAL.pdf
22. Goldberg Y., Ritov Y., Mandelbaum A. *The Best Linear Unbiased Estimator for Continuation of a Function.* Submitted to the Annals of Applied Statistics, 2010. http://arxiv.org/PS_cache/arxiv/pdf/1005/1005.1863v1.pdf
23. Zeltyn S., Mandelbaum A. *Service Engineering: Data-Based Course Development and Teaching.* To be published in INFORMS Transactions on Education, Vol. 11, No. 1, 2010; the special issue on "Teaching Service and Retail Operations Management", 2009. http://iew3.technion.ac.il/serveng/References/Gans-Koole-Mandelbaum- CCReview.pdf Full version: http://iew3.technion.ac.il/serveng/References/teaching_paper.pdf
24. The OCR Project (IBM, Rambam, Technion). *Services Research in Hospitals, or The Multi-Disciplinary Science of Patients Care*. Invited to the Journal of Service Research, June 2009, as the winner of the inaugural 2009 Service Science Innovation Partnership Award competition. http://iew3.technion.ac.il/serveng/References/Article_v11.pdf
25. Armony M., Mandelbaum A. *Routing and Staffing in Large-Scale Service Systems: The Case of Homogeneous Impatient Customers and Heterogeneous Servers*. Submitted for publication, 2009. http://iew3.technion.ac.il/serveng/References/IV_abandon_main.pdf, Technical appendix: http://iew3.technion.ac.il/serveng/References/IV_abandon_appendix.pdf
26. Khudyakov P., Feigin P.D., Mandelbaum A. *Designing a Call Center with an IVR (Interactive Voice Response).* Submitted for publication, 2009. http://iew3.technion.ac.il/serveng/References/Desighn_IVR.pdf
27. Aldor-Noiman S., Feigin P.D., Mandelbaum A. *Workload Forecasting for a Call Center: Methodology and a Case Study*. To be published in the Annals of Applied Statistics, 2009. http://iew3.technion.ac.il/serveng/References/PredictionMay1809.pdf
28. Marmor Y., Shtub A., Mandelbaum A., Wasserkrug S., Zeltyn S., Mesika Y., Greenshpan O., Carmeli B. *Toward Simulation-Based Real-Time Decision Support Systems for Emergency Departments*. Proceedings of the 2009 Winter Simulation Conference. M.D. Rossetti, R.R. Hill, B. Johansson, A. Dunkin and R.G. Ingalls, eds.
29. Gurvich I., Liberman P., Mandelbaum A. *Empirical Analysis of Skill Based Routing in Call Centers: A Queueing-Science Perspective.* 2009 MIT MSOM Conference.
30. Greenshpan O., Marmor Y., Wasserkrug S., Carmeli B., Vortman P., Basis F., Schwartz D., Mandelbaum A. *InEDvance*: *Advanced IT in Support of Emergency Department Management,* NGITS 2009, The 7th Conference on Next Generation Information Technologies and Systems.
31. Gurvich I., Armony M., Mandelbaum A. *Service Level Differentiation in Call Centers with Fully Flexible Servers.* Management Science, Vol. 54, No. 2, pp. 279–294, February 2008. http://iew3.technion.ac.il/serveng/References/Vdesign_pub.pdf Technical Appendix: http://iew3.technion.ac.il/serveng/References/Vdesign_ecompanion.pdf

32. Feldman Z., Mandelbaum A., Massey W.A., Whitt W. *Staffing of Time-Varying Queues to Achieve Time-Stable Performance***.** Management Science, Vol. 54, No.2, pp. 324–338, February 2008. http://iew3.technion.ac.il/serveng/References/Feldman2008pub.pdf, Internet supplement: http://iew3.technion.ac.il/serveng/References/Feldman2008e-companion.pdf, Unabridged version: http://iew3.technion.ac.il/serveng/References/Feldman_unabridged112705.pdf
33. Mandelbaum A., Zeltyn S. *Staffing Many-Server Queues with Impatient Customers: Constraint Satisfaction in Call Centers,* 2008. Under revision for Operations Research. http://iew3.technion.ac.il/serveng/References/MMNG_constraint_final.pdf; Internet Supplement: http://iew3.technion.ac.il/serveng/References/MMNG_constraint_supp_OR.pdf
34. Wasserkrug S., Tabu S., Zeltyn S., Gilat D., Lipets V., Feldman Z., Mandelbaum A. *Creating Operational Shift Scheduling for Third Level IT Support: Challenges, Models and Case Study*. Int. J. Services Operations and Informatics, Vol. 3, Nos. 3/4, 2008. http://iew3.technion.ac.il/serveng/References/Third%20Level%20Support%20IJTM1_cs.pdf http://iew3.technion.ac.il/serveng/References/Third%20Level%20SupportJapanese.pdf
35. Mandelbaum A., Zeltyn S. *Service Engineering in Action: The Palm/Erlang-A Queue, with Applications to Call Centers,* 2007. Advances in Services Innovations, pp. 17–48, Spath D., Fähnrich, K.-P. (Eds.), Springer-Verlag. http://iew3.technion.ac.il/serveng/References/MMNM_IAO.pdf
36. Aldor S., Feigin P.D. *Modeling Arrivals to the Agents Queue in a Call Center*, September, 2007.
37. Brown L., Gans N., Mandelbaum A., Sakov A., Zeltyn S., Zhao L., Haipeng S. *Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective*, JASA, Vol. 100, pp. 36–50, 2005. http://ie.technion.ac.il/serveng
38. Zeltyn S., Marmor Y., Mandelbaum A., Carmeli B., Greenshpan O., Mesika Y, Wasserkrug S., Vortman P., Schwartz D., Moskovitch K., Tzafrir S., Basis F., Shtub A., Lauterman T. *Simulation-Based Models of Emergency Departments: Real-Time Control, Operations Planning and Scenario Analysis*. Submitted to ACM Transactions on Modeling and Computer Simulation (TOMACS).
39. Mandelbaum A., Zeltyn S. *Service Engineering & Science: Simple Models at the Service of Complex Realities*, Service Science, Management, and Engineering (SSME): Education for the 21st Century. B. Hefley and W. Murphy. Eds., Springer, 2008.
40. Zeltyn S., Mandelbaum A. *Call Centers with Impatient Customers: Many-Server Asymptotics of the M/M/n+G Queue*, QUESTA, Vol. 51, Nos. 3/4, pp. 361–402, 2005. http://ie.technion.ac.il/serveng
41. Trofimov V., Feigin P., Mandelbaum M., Ishay E., Nadjarov E. *DATA-MOCCA: Data Model for Call Center Analysis*. Technion Report, Technion, May 2004 (56 pages). Last version: June 2006. http://ie.technion.ac.il/serveng
42. Marmor Y., Wasserkrug S., Carmeli B., Greenshpan O.,Vortman P., Schwartz D., Moskovitch K., Tzafrir S., Basis F., Shtub A., Lauterman T., Mandelbaum A. *RFID-Based Business Process Transformation: Value Assessment in Hospital Emergency Departments,* IBM + Rambam + Technion, OCR (Open Collaborative Research).
43. Mandelbaum A., Sakov A., Zeltyn S. *Empirical Analysis of a Call Center*, Technion Report, Technion, 2000 (73 pages). **Awarded the Students' Mitchner Prize for "Quality Sciences and Quality Management"**, Technion, 2001. Both text and overheads (by Anat Sakov). http://ie.technion.ac.il/serveng

Theses/Projects

44. Khudyakov P. *Statistical Analyses of Call Center Data.* Ph.D. Thesis, August 2010. http://iew3.technion.ac.il/serveng/References/thesis_final_el.pdf
45. Yom-Tov G. *Queues in Hospitals: Queueing Networks with ReEntering Customers in the QED Regime*. Ph.D. Thesis, Technion, June 2010. http://iew3.technion.ac.il/serveng/References/galit_PHD_submitted.pdf http://iew3.technion.ac.il/serveng/References/Seminar2010.pdf
46. Marmor Y. *Emergency-Departments Simulation in Support of Service-Engineering: Staffing, Design, and Real-Time Tracking*. Ph.D. Thesis, Technion, February 2010. http://iew3.technion.ac.il/serveng/References/Yariv_PHD.pdf http://iew3.technion.ac.il/serveng/References/YarivSeminar_n.pdf

47. Maman S. *Uncertainty in the Demand for Service: The Case of Call Centers and Emergency Departments.* M.Sc. Thesis, Technion, April 2009.
http://iew3.technion.ac.il/serveng/References/Thesis_Shimrit.pdf
http://iew3.technion.ac.il/serveng/References/Seminar_Shimrit.pdf
http://iew3.technion.ac.il/serveng/References/Thesis_Shimrit_Heb.pdf
48. Tseytlin Y. *Queueing Systems with Heterogeneous Servers: On Fair Routing of Patients in Emergency Departments*. M.Sc. Thesis, Technion, April 2009.
http://iew3.technion.ac.il/serveng/References/thesis-yulia.pdf
http://iew3.technion.ac.il/serveng/References/seminar_yulia_9_2.pdf
http://iew3.technion.ac.il/serveng/References/heb_summary_yulia.pdf
49. Feldman Z. *Optimal Staffing of Systems with Skills-Based-Routing (SBR)*. M.Sc. Thesis, Technion, 2008. http://iew3.technion.ac.il/serveng/References/Zohar_Seminar.ppt
http://iew3.technion.ac.il/serveng/References/Zohar_Thesis.pdf
50. Tseytlin Y., Zviran A. *Simulation of Patients Routing from an Emergency Department to Internal Wards in Rambam Hospital.* Technion, November 2008.
http://iew3.technion.ac.il/serveng/References/project_yulia_asaf.pdf
http://iew3.technion.ac.il/serveng/References/presentation_yulia_asaf.pdf
51. Shaikhet G. *Control of Many-Servers Queueing Systems in Heavy Traffic*. Ph.D. Thesis, Technion, 2007. http://iew3.technion.ac.il/serveng/References/PhD_lecture_gennady.pdf
http://iew3.technion.ac.il/serveng/References/PhD_gennady.pdf
52. Rozenshmidt L. *On Priority Queues with Impatient Customers: Stationary and Time-Varying Analysis.* M.Sc. Thesis, Technion, 2007.
http://iew3.technion.ac.il/serveng/References/Luba_seminar.ppt
http://iew3.technion.ac.il/serveng/References/thesis_Luba_Eng.pdf
http://iew3.technion.ac.il/serveng/References/thesis_Luba_Heb.pdf
53. Khudyakov P. *Designing a Call Center with an IVR (Interactive Voice Response).* M.Sc. Thesis, Technion, 2006. http://iew3.technion.ac.il/serveng/References/Polina_presentation_16.ppt
http://iew3.technion.ac.il/serveng/References/thesis_polyna.pdf
54. Aldor-Noiman S. *Forecasting Demand for a Telephone Call Center: Analysis of Desired versus Attainable Precision*. M.Sc. Thesis, Technion, 2006.
http://iew3.technion.ac.il/serveng/References/Sivan_Seminar.ppt
http://iew3.technion.ac.il/serveng/References/SivanThesisFinal.pdf
55. Zeltyn S. *Call Centers with Impatient Customers: Exact Analysis and Many-Server Asymptotics of the M/M/n+G queue*. PhD Thesis, Technion, 2005.
http://iew3.technion.ac.il/serveng/References/MMNG_seminar.pdf
http://iew3.technion.ac.il/serveng/References/MMNG_thesis.pdf
56. Feldman Z. *Staffing of Time-Varying Queues to Achieve Time-Stable Performance*. Technion, June 2004. http://iew3.technion.ac.il/serveng/References/Zohar_Project.pdf
57. Gurvich I. *Design and Control of the M/M/N Queue with Multi-Class Customers and Many Servers.* M.Sc. Thesis, Technion, July 2004.
http://iew3.technion.ac.il/serveng/References/Gurvich_Seminar.pdf
http://iew3.technion.ac.il/serveng/References/ThesisItay.pdf
58. Issaev E. *Fitting Phase-Type Distributions to Data from a Telephone Call Center.* M.Sc. Thesis, Technion, March 2003. http://iew3.technion.ac.il/serveng/References/Seminar_presentation.ppt
http://iew3.technion.ac.il/serveng/References/Thesis.pdf
59. Marmor Y. *Developing a Simulation Tool for Analyzing Emergency Department Performance*. M.Sc. Thesis, Technion, August 2003.
http://iew3.technion.ac.il/serveng/References/thesis_Yariv_Eng.pdf
http://iew3.technion.ac.il/serveng/References/thesis_Yariv_Heb.pdf

Graduate Student Proposals

60. Zaied I. *The Offered Load in Fork-Join Networks: Calculations and Applications to Service Engineering of Emergency Department*. M.Sc. Research Proposal, Technion, March 2010.
http://iew3.technion.ac.il/serveng/References/Itamar%20Zaied's%20Research%20Proposal.pdf
61. Senderovich A. *Call Centers with Skills-Based-Routing: Simulation-Based State-Inference and Dynamic-Staffing.* M.Sc. Research Proposal, Technion, February 2010.

http://iew3.technion.ac.il/serveng/References/Arik_Research_Proposal_Final.pdf

62. Zviran A. *Fork-Join Networks in Heavy Traffic: Diffusion Approximations and Control*. M.Sc. Research Proposal, Technion, August 2008. http://iew3.technion.ac.il/serveng/References/Asaf_research_proposal.pdf
63. Reich M. *The Workload Process: Modelling, Inference and Applications*. M.Sc. Research Proposal, Technion, December 2007. http://iew3.technion.ac.il/serveng/References/proposal_Michael.pdf

Graduate Students Projects (see also OCR theses below)

64. Zaied I. (supervised jointly with H. Kaspi): *The Offered Load of Fork-Join Networks, with Applications to Health Care.* M.Sc. Technion, 2009–.
65. Zviran A. (supervised jointly with R. Atar): *On Fork-Join (Split-Match) Networks*. M.Sc., Technion, 2010.
66. Tseytlin Y. *Queueing Systems with Heterogeneous Servers: On Fair Routing of Patients in Emergency Departments.* M.Sc., Technion, 2009 **(Summa Cum Laude)**.
67. Maman S. (supervised jointly with S. Zeltyn): *Uncertainty in the Demand for Service: The Case of Call Centers and Emergency Departments.* M.Sc. in Statistics, 2009. **Winner of the 2009 ORSIS Mechraz Prize for Best Graduate Thesis in Operations Research.**
68. Rosenschmidt L. *On Queues with Impatient Customers: Theory, Approximation and Simulation.* M.Sc., Technion, 2008.
69. Feldman Z. *Optimal Staffing of Systems with Skills-Based-Routing*. M.Sc., Technion, 2009 **(Summa Cum Laude)**.
70. Marmor Y. *Emergency-Departments Simulation in Support of Service-Engineering: Staffing, Design, and Real-Time Tracking*. Ph.D., Technion, 2010.
71. Yom-Tov G. *Performance Analysis of Health-Care Systems*. Ph.D., Technion, 2010.
72. Reich M. (supervised jointly with Y. Ritov): *The Workload Process: Modelling, Analysis and Inference*. M.Sc., Technion, 2010.
73. Aldor S. (main advisor: P. Feigin): *Statistical Inference of Call Center Data*. M.Sc., Technion, 2006.

Student Projects (see also OCR Student projects below)

74. Plonsky O. (supervised jointly with S. Zeltyn): *Validating Models of Call Centers*, 2010.
75. Koren A. *On Workload and Offered-Load Programming*, 2010.
76. Paldi R., Hecht A., Israeli A., Solo M.; Kutsy K. (SEE Advisor): *On Fusing Financial and Operational Data*, Mizrahi Bank (Cohen A.), 1 year, 2010.
77. Gigi L., Koshman K., Rath N., Saks Y.; Khudiakov P. (Ph.D. advisor): *Retail Services at Bank Hapoalim* (Shlasky T., Sol E.), 1 year, 2008–9.
78. Babajani N., Hod A., Liberman P. *SBR at Pelephone* (Shiri Y., Ater E.), 1 year, 2008.
79. Kagan Y., Katz O., Moses L., Khudiakov P. (Ph.D. advisor), *Mortgage Tele-services at Bank Hapoalim* (Maklef G., Hugi Z., Aviad E.), 1 year, 2007–2008.

OCR Research

Research papers

80. *Services Research in Hospitals, or The Multi-Disciplinary Science of Patients Care,* Marmor Y. , Shtub A., Mandelbaum A., Vasserkrug S., Carmeli B., Vortman P., Schwarts D., Basis F., Assaf J. Gopher D., Auarbach-Shpak Y., Paper for JS (Journal of Service Research).
81. *Toward Simulation-Based Real-Time Decision Support Systems For Emergency Departments,* Marmor Y. , Shtub A., Mandelbaum A., Vasserkrug S., Zeltyn S., Carmeli B., Greenshpan O., Masika Y., Paper for WSC 2009 Conference.
82. *InEDvance: Advanced IT in Support of Emergency Department Management,* Marmor Y. , Mandelbaum A., Vasserkrug S., Carmeli B., Greenshpan O., Vortman P., Schwartz D., Basis F., Paper for NGITS 2009 Conference.

83. *Simulation-based Models of Emergency Departments: Real-Time Control, Operations Planning and Scenario Analysis,* Zeltyn S., Marmor Y., Mandelbaum A., Carmeli B., Greenshpan O., Mesika Y., Wasserkrug S., Vortman P., Schwartz D., Moskovitch K., Tzafrir S., Basis F., Shtub A., Lauterman T., Paper for TOMACS.

Graduate works (PhD, MSc)

84. *Queues in Hospitals: Semi-Open Queueing Networks in the QED Regime*, Yom-Tov G.**,** Ph.D.
85. *The Workload Process: Modelling, Inference and Applications*, Reich M.**,** M.Sc.
86. *Uncertainty in the Demand for Service: The Case of Call Centers and Emergency Departments*, Maman S.**,** M.Sc.
87. *Queueing Systems with Heterogeneous Servers: Improving Patients' Flow in Hospitals***,** Tseytlin Y., M.Sc.
88. *Emergency-Departments Simulation in Support of Service-Engineering: Staffing, Design, and Real-Time Tracking***,** Marmor Y., Ph.D.

Student projects

89. Plonsky O., David N., Dourban A., Gologorsky M.; Zaied I. (M.Sc. Advisor), Efrat D. (Ph.D. Advisor): *Fairness in the Allocation of Work among a Hospital's Maternity Wards, Rambam Hospital* (Kranzler M., Tzafrir S.), 2010.
90. Koren A., Yungelson A., Shvered T., Stiro Z.; Marmor Y. (Ph.D. Advisor), Zychlinski N.(M.Sc. Advisor): *RFID-Based Control of a Mass-Casualty-Event* (Dr. Israelit S.), 1 year, 2010.
91. Horev A., Perez L., Goldshtein A., Hamo A.*;* Marmor Y. (Ph.D. advisor): *Comparison of Four Possible Operational Models for ED* (Dr. Schwartz D.)*,* 2009.
92. Oppenheim Y., Raanan N., Ston B., Arieli Y.; Marmor Y. (Ph.D. advisor): *Designing an RFID System in an Emergency Department, Rambam Hospital* (Dr. Freed T., Dr. Schwartz D.), 1 year, 2009.
93. Kubi Y., Shimon S., Yuviler N., *Choosing the Most Effective Operational Model for the ED* (Dr. Israelit S.), 2008.
94. Kubi Y., Shimon S., Yuviler N.; Luzon Y. (Ph.D. advisor): *Analysis of the Pre- Operative Process at the Rambam Hospital* (Levy D.), 1 year, 2008.
95. Natur Tawheed, Valin V.; Marmov Y. (Ph.D. advisor): *Architectures of Emergency Departments—Comparative Analysis via DEA*, 1 year, 2008.
96. Tseytlin Y., Zviran A. S*imulation of Patients Routing from an Emergency Department to Internal Wards in Rambam Hospital*, 2008.
97. Badran K., Haas I., Huli Rasha; Marmor Y. (Ph.D. advisor): *Operational Aspects of the Transition to a Permanent ED at Rambam Hospital* (Dr. Fuad Basis, Dr. Michelson M.), 1/2 year, 2008.
98. Elkin K., Rosenberg N., Levy S., Shiri Y. *Patients' Flow from the Emergency Department to the Internal Wards at the Rambam Hospital* (Dr. Barel Y., Shiloah M.), 1 year, September 2007 (Supplement: A Business Plan for RFID in the ED).

Technion Publications—Using and Acknowledging SEE Data

1. Mandelbaum A., Feldman Z. *Using Simulation-Based Stochastic Approximation to Optimize Staffing of Systems with Skills-Based Routing*. Proceedings of the 2010 Winter Simulation Conference, December 2010.
   http://iew3.technion.ac.il/serveng/References/wscSAStaffing_v3.pdf

Abstract: In this paper, we consider the problem of minimizing the operational costs of systems with Skills-Based-Routing (SBR). In such systems, customers of multiple classes are routed to servers of multiple skills. In the settings we consider, each server skill is associated with a corresponding cost, and service level can either appear as a strong constraint or incur a cost. The solution we propose is based on the Stochastic Approximation (SA) approach. Since SBR models are analytically intractable in general, we use computer simulation to evaluate service-level measures. Under the assumption of convexity of the service-level as functions in staffing levels, SA provides an analytical proof of convergence, together with a rate of convergence. We show, via numerical examples, that although the convexity assumption does not hold for all cases and all types of service-level objectives, the algorithm nevertheless identifies the optimal solution.

2. Gans N., Liu N., Mandelbaum A., Shen H., Ye H. *Service Times in Call Centers: Agent Heterogeneity and Learning with some Operational Consequences*. To appear in a book on the occasion of Larry Brown's 70th birthday, December 2010
   http://iew3.technion.ac.il/serveng/References/agent-heterogeneity-Brown-Book-FINAL.pdf

Abstract: Telephone call centers are data-rich environments that, until recently, have not received sustained attention from academics. For about a decade now, we are fortunate to have worked with our colleague, mentor and friend, Larry Brown, on the collection and analysis of large call-center datasets. This work has provided many fascinating windows into the world of call-center operations, stimulating further research and affecting management practice. Larry's inexhaustible curiosity and creativity, sharp insight and unique technical power, have continuously been an inspiration to us. We look forward to collaborating and learning from him on many occasions to come. In this paper, we study operational heterogeneity of call center agents. Our proxy for heterogeneity is agents' service times (call durations), a performance measure that prevalently enjoys tight management control. Indeed, managers of large call centers argue that a 1-second increase/decrease in average service time can translate into additional/reduced operating costs on the order of millions of dollars per year. We are motivated by an empirical analysis of call-center data, which identifies both short- term and long-term factors associated with agent heterogeneity. Operational consequences of such heterogeneity are then illustrated via discrete event simulation. This highlights the potential benefits of analyzing individual agents' operational histories. We are thus naturally led to a detailed analysis of agents' learning-curves, which reveals various learning patterns and opens up new research opportunities.

3. Mandelbaum A., Momcilovic P., Tseytlin Y. *On Fair Routing From Emergency Department to Hospital Wards: QED Queues with Heterogeneous Servers*. Submitted to Management Science, 2010. http://iew3.technion.ac.il/serveng/References/fr6.pdf

Abstract: The process of patients' routing from an Emergency Department (ED) to Internal Wards (IWs) is often a hospital's bottleneck. Motivated by this process in Anonymous Hospital, we analyze queueing systems with heterogeneous server pools, where the pools represent the wards and servers are beds. Our queueing system, with a single centralized queue and several server pools, forms an inverted-V model. We introduce the Randomized Most-Idle (RMI) routing policy and analyze it in the QED (Quality and Efficiency Driven) regime, which is natural in our setting. The RMI policy results in the same server fairness (measured by idleness ratios) as the Longest-Idle Server First (LISF) policy, which is commonly used in call centers and considered fair. However, RMI utilizes

only the information on the number of idle servers in different pools while LISF requires information that is unavailable in hospitals on a real-time basis.

4. Goldberg, Y., Ritov, Y., Mandelbaum, A. *The Best Linear Unbiased Estimator for Continuation of a Function*. Submitted to the Annals of Applied Statistics, 2010. http://arxiv.org/PS_cache/arxiv/pdf/1005/1005.1863v1.pdf

Abstract: We show how to construct the best linear unbiased predictor (BLUP) for the continuation of a curve in a spline-function model. We assume that the entire curve is drawn from some smooth random process and that the curve is given up to some cut point. We demonstrate how to compute the BLUP efficiently. Confidence bands for the BLUP are discussed. Finally, we apply the proposed BLUP to real-world call center data. Specifically, we forecast the continuation of both the call arrival counts and the workload process at the call center of a commercial bank.

5. Atar, R. *A Diffusion Regime with Non-degenerate Slowdown*, June 2010. http://webee.technion.ac.il/people/atar/NDS-rev.pdf

Abstract: We study a diffusion regime, earlier considered by Gurvich, Mandelbaum, Shaikhet and Whitt in the case of the M/M/N queue, that is, in a sense that we make precise, a midpoint between two well-known heavy traffic diffusion regimes, the conventional and the quality and efficiency driven regimes. Unlike the other two, this regime, that we call the non-degenerate slowdown regime, enjoys the property that delay and service time are of the same order of magnitude, a property that is often desirable from a modeling viewpoint. Our main result is that in the case of heterogeneous exponential multi-server systems, this regime gives rise to new limit processes for the sojourn time. In particular, the joint limit law of the delay and service time processes is identified as a reflected Brownian motion and an independent process, whose marginal is a size-biased mixture of exponentials. Our results also motivate the formulation and study of new diffusion control problems, based on sojourn time cost.

6. Khudyakov P., Feigin P.D., Mandelbaum A. *Designing a Call Center with an IVR (Interactive Voice Response)*. Submitted for publication, 2009. http://iew3.technion.ac.il/serveng/References/Desighn_IVR.pdf

Abstract: A call center is a popular term for a service operation that caters to customers' needs via the telephone. A call center typically consists of agents that serve customers, telephone lines, an Interactive Voice Response (IVR) unit, and a switch that routes calls to agents.
In this paper we study a Markovian model for a call center with an IVR. We calculate operational performance measures, such as the probability for a busy signal and the average wait time for an agent. Exact calculations of these measures are cumbersome and they lack insight. We thus approximate the measures in an asymptotic regime known as QED (Quality & Efficiency Driven) or the Halfin-Whitt regime, which accommodates moderate to large call centers. The approximations are both insightful and easy to apply (for up to 1000's of agents). They yield, as special cases, known and novel approximations for the M/M/N/N (Erlang-B), M/M/S (Erlang-C) and M/M/S/N queue.

7. Aldor-Noiman S., Feigin P.D., Mandelbaum A. *Workload Forecasting for a Call Center: Methodology and a Case Study*. To be published in the Annals of Applied Statistics. 2009. http://iew3.technion.ac.il/serveng/References/PredictionMay1809.pdf

Abstract: Today's call center managers face multiple operational decision-making tasks. One of the most common is determining the weekly staffing levels to ensure customer satisfaction and meeting their needs while minimizing service costs. An initial step for producing the weekly schedule is forecasting the future system loads which involve predicting both arrival counts and average service times. We introduce an arrival count model which is based on a

mixed Poisson process approach. The model is applied to data from an Israeli Telecom company call center. In our model, we also consider the effect of events such as billing on the arrival process and we demonstrate how to incorporate them as exogenous variables in the model. After obtaining the forecasted system load, in large call centers, a manager can choose to apply the QED (Quality-Efficiency Driven) regime's "square root staffing" rule in order to balance the offered-load per server with the quality of service. Implementing this staffing rule requires that the forecasted values of the arrival counts and average service times maintain certain levels of precision. We develop different goodness of fit criteria that help determine our model's practical performance under the QED regime. These show that during most hours of the day the model can reach desired precision levels.

8. Gurvich I., Liberman P., Mandelbaum A. *Empirical Analysis of Skill Based Routing in Call Centers: A Queueing-Science Perspective.* 2009 MIT MSOM Conference.

9. Brown L., Gans N., Mandelbaum A., Sakov A., Zeltyn S., Zhao L., Haipeng S. *Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective*, JASA*,* 100, 36–50, 2005. http://iew3.technion.ac.il/serveng/References/JASA_callcenter.pdf

Abstract: A call center is a service network in which agents provide telephone-based services. Customers who seek these services are delayed in tele-queues. This article summarizes an analysis of a unique record of call center operations. The data comprise a complete operational history of a small banking call center, call by call, over a full year. Taking the perspective of queueing theory, we decompose the service process into three fundamental components: arrivals, customer patience, and service durations. Each component involves different basic mathematical structures and requires a different style of statistical analysis. Some of the key empirical results are sketched, along with descriptions of the varied techniques required. Several statistical techniques are developed for analysis of the basic components. One of these techniques is a test that a point process is a Poisson process. Another involves estimation of the mean function in a nonparametric regression with lognormal errors. A new graphical technique is introduced for nonparametric hazard rate estimation with censored data. Models are developed and implemented for forecasting of Poisson arrival rates. Finally, the article surveys how the characteristics deduced from the statistical analyses form the building blocks for theoretically interesting and practically useful mathematical models for call center operations.

10. Zeltyn S., Mandelbaum A. *Call Centers with Impatient Customers: Many-Server Asymptotics of the M/M/n+G Queue*, QUESTA, 51 (3/4), 361–402, 2005. http://iew3.technion.ac.il/serveng/References/regimes.pdf

Abstract: The subject of the present research is the M/M/*n*+G queue. This queue is characterized by Poisson arrivals at rate λ, exponential service times at rate μ, *n* service agents and generally distributed patience times of customers. The model is applied in the call center environment, as it captures the tradeoff between operational efficiency (staffing cost) and service quality (accessibility of agents). In our research, three asymptotic operational regimes for medium to large call centers are studied. These regimes correspond to the following three staffing rules, as λ and *n* increase indefinitely and μ held fixed:
Efficiency-Driven (ED): $n \approx (\lambda/\mu) \cdot (1 - \gamma), \quad \gamma > 0,$
Quality-Driven (QD): $n \approx (\lambda/\mu) \cdot (1 + \gamma), \quad \gamma > 0,$ and
Quality and Efficiency Driven (QED): $n \approx \lambda/\mu + \beta\sqrt{\lambda/\mu}, \quad -\infty < \beta < \infty.$

In the ED regime, the probability to abandon and average wait converge to constants. In the QD regime, we observe a very high service level at the cost of possible overstaffing. Finally, the QED regime carefully balances quality and efficiency: agents are highly utilized, but the probability to abandon and the average wait are small (converge to zero at rate $1/\sqrt{n}$). Numerical experiments demonstrate that, for a wide set of system parameters, the QED formulae provide excellent approximation for exact M/M/*n*+G performance measures. The

much simpler ED approximations are still very useful for overloaded queueing systems. Finally, empirical findings have demonstrated a robust linear relation between the fraction abandoning and average wait. We validate this relation, asymptotically, in the QED and QD regimes.

11. Mandelbaum A., Sakov A., Zeltyn S. *Empirical Analysis of a Call Center*, Technion Report, Technion, 2000 (73 pages). Awarded the Students' Mitchner Prize for "Quality Sciences and Quality Management", Technion, 2001. http://iew3.technion.ac.il/serveng/References/ccdata.pdf.

Call center is the common term for describing a telephone-based human-service operation. A call center provides tele-services, namely services in which the customers and the service agents are remote from each other. The agents, who sit in cubicles, constitute the physical embodiment of the call center: with numbers varying from very few to many hundreds, they serve customers over the phone, while facing a computer terminal that outputs and inputs customer data. The customers, who are only virtually present, are either being served, or they are waiting in, what we call, tele-queues: up to possibly thousands of customers sharing a phantom queue, invisible to each other and the agents serving them, waiting and accumulating impatience until one of two things happens: an agent is allocated to serve them (through a supporting software), or they abandon the tele-queue, plausibly due to impatience that has built up to exceed their anticipated worth of the service. The world of call centers is vast: some estimate that 70% of all customer-business interactions occur in call centers; that $700 billions in goods and services were sold through call centers in 1997, and this figure has been expanding 20% annually; and that 3% of the U.S. working population is currently employed in call centers. (This amounts to 1.55 million agents, and some estimates actually go up to 6 million). The leading-edge call center is a complex socio-technical system: its hundreds of agents could cater to thousands of customers per hour, in a way that the average wait is measured in a few seconds and agents' utilization exceeds 90%. Such simultaneous attainment of superb service quality with extreme resource efficiency is achievable, despite ample stochastic variability, through scale-economies of unparalleled magnitudes; and all this is possible only in the unique frictionless environment of computer-telephony integration and automatic call distribution. Some view call centers as the business frontiers and others as the sweat-shops of the 21st century. Either way, call centers provide ample uncharted challenges for researchers in multi-disciplines, from the soft (e.g. Psychology, Sociology), through functional management (e.g. Marketing, Information Systems), to the exact (e.g. Computer Science, Mathematics). One should note that the challenges are, in fact, expanding: there exist an increasing number of multi-media call centers that can provide, in addition to the telephone, also video, Internet, fax and e-mail services. (The term customer contact center has been used to accommodate this broader connotation of a tele-service.) Our paper targets primarily researchers in Statistics, Operations-Research (especially Queueing Theory, and even more so Queueing Science, Operations Management, and Industrial Engineering. We believe that it is also of interest to researchers in telecommunications, and of use to managers that either run or oversee the operations of medium to large call centers.

Out-of-Technion Publications—Using and Acknowledging SEE Data

12. K. Kearton. *Correlating Temporal Rules to Time-Series Data with Rule-Based Intuition*, March 2010. http://edocs.nps.edu/npspubs/scholarly/theses/2010/Mar/10Mar_Kearton.pdf

Abstract: Analysts are frequently confronted with time-series data. A simple form is magnitude (or count) and time frame, whether the data is number of e-mails sent, number of cell phones called, purchases made by volume or cost, or a variety of other time-derived data. Studying the temporal dimension of data allows analysts more opportunities to find relational ties and trends in data, classify or group like activity, and even help narrow the search space of massively complex and large datasets. This thesis presents a new approach called the Rule Based Intuition (RBI) system that can evaluate time-series data by finding the best fitting rule, from a repository of known rules, to quickly infer information about the data. This approach is most applicable for analysts viewing large sets of data who wish to classify or correlate data from users' temporal activity.

13. A. M. De Livera, R. J. Hyndman. *Forecasting Time Series with Complex Seasonal Patterns using Exponential Smoothing*, December 2009. http://robjhyndman.com/papers/ComplexSeasonality.pdf

Abstract: A new innovations state space modeling framework, incorporating Box-Cox transformations, Fourier series with time varying coefficients and ARMA error correction, is introduced for forecasting complex seasonal time series that cannot be handled using existing forecasting models. Such complex time series include time series with multiple seasonal periods, high frequency seasonality, non-integer seasonality and dual-calendar effects. Our new modelling framework provides an alternative to existing exponential smoothing models, and is shown to have many advantages. The methods for initialization and estimation, including likelihood evaluation, are presented, and analytical expressions for point forecasts and interval predictions under the assumption of Gaussian errors are derived, leading to a simple, comprehensible approach to forecasting complex seasonal time series .Our trigonometric formulation is also presented as a means of decomposing complex seasonal time series, which cannot be decomposed using any of the existing decomposition methods. The approach is useful in a broad range of applications, and we illustrate its versatility in three empirical studies where it demonstrates excellent forecasting performance over a range of prediction horizons. In addition, we show that our trigonometric decomposition leads to the identification and extraction of seasonal components, which are otherwise not apparent in the time series plot itself.

14. D. Bestimas, X. Vinh Doan. *Robust and Data-Driven Approaches to Call Centers*, November 2008 http://web.mit.edu/dbertsim/www/papers/Robust%20Optimization/DataDriven%20and%20Robust%20Optimization%20Approaches%20to%20Call%20Centers.pdf

Abstract: We propose both robust and data-driven approaches to a fluid model of call centers that incorporates random arrival rates with abandonment to determine staff levels and dynamic routing policies. Resulting models are tested with real data obtained from the call center of a US bank. Computational results show that the robust fluid model is significantly more tractable as compared to the data-driven one and producing better solutions to call centers in most experiments."

15. H. Shen, J. Z. Huang. *Interday Forecasting and Intraday Updating of Call Center Arrivals,* Manufacturing & Service Operations Management, Vol. 10, No. 3, summer 2008, pp. 391–410. http://msom.journal.informs.org/cgi/content/abstract/10/3/391

Abstract: Accurate forecasting of call arrivals is critical for staffing and scheduling of a telephone call center. We develop methods for interday and dynamic intraday forecasting of incoming call volumes. Our approach is to treat the intraday call volume profiles as a high-dimensional vector time series. We propose first to reduce the dimensionality by singular value decomposition of the matrix of historical intraday profiles and then to apply time series and regression techniques. Our approach takes into account both interday (or day-to-day) dynamics and intraday (or within-day) patterns of call arrivals. Distributional forecasts are also developed. The proposed methods are data driven, appear to be robust against model assumptions in our simulation studies, and are shown to be very competitive in out-of-sample forecast comparisons using two real data sets. Our methods are computationally fast; it is therefore feasible to use them for real-time dynamic forecasting."

16. J.W. Taylor. *A Comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center*, Management Science, February 2008, Vol. 54, pp. 253–265. http://users.ox.ac.uk/~mast0315/CallCenterFcstComparison.pdf

Abstract: Predictions of call center arrivals are a key input to staff scheduling models. It is, therefore, surprising that simplistic forecasting methods dominate practice, and that the research literature on forecasting arrivals is so small. In this paper, we evaluate univariate time series methods for forecasting intraday arrivals for lead times from one half-hour ahead to two weeks ahead. We analyze five series of intraday arrivals for call centers operated by a retail bank in the UK. A notable feature of these series is the presence of both an intraweek and an intraday seasonal cycle. The methods considered include seasonal ARIMA modeling; periodic AR modeling; an extension of Holt-Winters exponential smoothing for the case of two seasonal cycles; robust exponential smoothing based on exponentially weighted least absolute deviations regression; and dynamic harmonic regression, which is a form of unobserved component state space modeling. Our results indicate strong potential for the use of seasonal ARIMA modeling and the extension of Holt-Winters for predicting up to about two to three days ahead and that, for longer lead times, a simplistic historical average is difficult to beat. We find a similar ranking of methods for call center data from an Israeli bank.

17. M. Concepcion Ausin, Michael P. Wiper, Rosa E. Lillo. *Bayesian Prediction of the Transient Behaviour and Busy Period in Short and Long-Tailed GI/G/1 Queueing Systems*, January 2008. http://www.udc.es/dep/mate/Dpto_Matematicas/Investigacion/ie_publicacion/Ausin.Wiper.Lillo.pdf

Abstract: In this paper, we describe how to carry out Bayesian inference for the transient behaviour and busy period in a single server queueing system with general and unknown distributions for the interarrival and service times. Both interarrival and service time distributions are approximated using the dense family of Coxian distributions. A suitable reparameterization enables us to define a non-informative prior and implement Bayesian inference for the Coxian model using reversible jump Markov chain Monte Carlo methods. An advantage of our procedure is that we are well able to approximate heavy tailed interarrival and service time distributions such as the Pareto. Our procedure for estimating the system measures is based in recent theoretical results for the Coxian/Coxian/1 system. We propose a numerical technique that can be performed for every MCMC iteration which allows us to estimate interesting measures such as the transient queue length and waiting time distributions and the duration of a busy period. We illustrate our approach with both simulated and real data.

18. Anders Gorst-Rasmussen, Martin B. Hansen. *Asymptotic Inference for Waiting Times and Patiences in Queues with Abandonment*, July 2007. http://www.math.aau.dk/fileadmin/user_upload/www.math.aau.dk/Forskning/Rapportserien/R-2007-14.pdf

Abstract: Motivated by applications in call center management, we propose a framework based

on empirical process techniques for inference about the waiting time and patience distribution in multiserver queues with abandonment. The framework rigorises heuristics based on survival analysis of independent and identically distributed observations by allowing correlated successive waiting times. Assuming a regenerative structure of the sequence of offered waiting times, we establish asymptotic properties of estimators of limiting distribution functions and derived functionals. We discuss construction of bootstrap confidence intervals and statistical tests, including a simple bootstrap two-sample test for comparing patience distributions. The methods are exemplified in a small simulation study, and a real data example is given involving comparison of patience distributions for two customer classes in a call center.

19. A. Thümmler, P. Buchholz, M. Telek. *A Novel Approach for Phase-Type Fitting with the EM Algorithm*, July 2006. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1673383

Abstract: The representation of general distributions or measured data by phase-type distributions is an important and non-trivial task in analytical modeling. Although a large number of different methods for fitting parameters of phase-type distributions to data traces exist, many approaches lack efficiency and numerical stability. In this paper, a novel approach is presented that fits a restricted class of phase-type distributions, namely mixtures of Erlang distributions, to trace data. For the parameter fitting an algorithm of the expectation maximization type is developed. The paper shows that these choices result in a very efficient and numerically stable approach which yields phase-type approximations for a wide range of data traces that are as good or better than approximations computed with other less efficient and less stable fitting methods. To illustrate the effectiveness of the proposed fitting algorithm, we present comparative results for our approach and two other methods using six benchmark traces and two real traffic traces as well as quantitative results from queueing analysis.

20. J. Weinberg, L.D. Brown, J.R. Stroud. *Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data*, June 2006.

Abstract: A call center is a centralized hub where customer and other telephone calls are dealt with by an organization. In today's economy, they have become the primary point of contact between customers and businesses. Accurate prediction of the call arrival rate is therefore indispensable for call center practitioners to staff their call center efficiently and cost effectively. This article proposes a multiplicative model for modeling and forecasting within-day arrival rates to a US commercial bank's call center. Markov chain Monte Carlo sampling methods are used to estimate both latent states and model parameters. One-day-ahead density forecasts for the rates and counts are provided. The calibration of these predictive distributions is evaluated through probability integral transforms. Furthermore, we provide one-day-ahead forecasts comparisons with classical statistical models. Our predictions show significant improvements of up to 25% over these standards. A sequential Monte Carlo algorithm is also proposed for sequential estimation and forecasts of the model parameters and rates.

21. J. Z. Huang, N. Liu, M. Pourahmadi, L. Liu. *Covariance Matrix Selection and Estimation via Penalized Normal Likelihood*, Biometrika (March 2006), Vol. 93, No. 1, pp. 85–98. http://biomet.oxfordjournals.org/content/93/1/85.full.pdf

Abstract: We propose a nonparametric method for identifying parsimony and for producing a statistically efficient estimator of a large covariance matrix. We reparameterise a covariance matrix through the modified Cholesky decomposition of its inverse or the one-step-ahead predictive representation of the vector of responses and reduce the nonintuitive task of modelling covariance matrices to the familiar task of model selection and estimation for a sequence of regression models. The Cholesky factor containing these regression coefficients is likely to have many off-diagonal elements that are zero or close to zero. Penalised normal likelihoods in this situation with L1 and L2 penalties are shown to be closely related to Tibshirani's (1996) LASSO approach and to ridge regression. Adding either penalty to the

likelihood helps to produce more stable estimators by introducing shrinkage to the elements in the Cholesky factor, while, because of its singularity, the L1 penalty will set some elements to zero and produce interpretable models. An algorithm is developed for computing the estimator and selecting the tuning parameter. The proposed maximum penalised likelihood estimator is illustrated using simulation and a real dataset involving estimation of a 102 x 102 covariance matrix.

22. H. Shen, L.D. Brown. *Nonparametric Modeling of Time-varying Customer Service Times at a Bank Call Center*, January 2006.
https://mercury.smu.edu.sg/rsrchpubupload/7292/c2006-1-1.pdf

Abstract: Call centers becoming increasingly important in our modern commerce. We are interested in modeling the time-varying pattern of average customer service times at a bank call center. Understating such a pattern is essential for efficient operation of a call center. The call service times are show to be lognormally distributed. Motivated by this observation and the important application, we propose a new method for inference about nonparametric regression curves when the errors are lognormally distributed. Estimates and pointwise confidence bands are developed. The method builds upon the special relationship between the lognormal distribution and the normal distribution, and improves upon a naïve estimation procedure that ignores this distributional structure. Our approach includes local nonparametric estimation for both the mean function and the *heteroscedastic* variance function of the logged data, and uses local polynomial regression as a fitting tool. A simulation study is performed to illustrate the method. We then apply the method to model the time-varying patterns of mean service times for different types of customer calls. Several operationally interesting finding are obtained and discussed.

23. Sandjai Bhulai, Wing Hong Kan, Elena Marchiori. *Nearest Neighbour Algorithms for Forecasting Call Arrivals in Call Centers*, August 2005.
http://www.cs.ru.nl/~elenam/WS2005-12.pdf

Abstract: In this paper we study a nearest neighbour algorithm for forecasting call arrivals to call centers. The algorithm does not require an underlying model for the arrival rates and it can be applied to historical data without pre-processing it. We show that this class of algorithms provides a more accurate forecast when compared to the conventional method that simply takes averages. The nearest neighbour algorithm with the Pearson correlation distance function is also able to take correlation structures, that are usually found in call center data, into account. Numerical experiments show that this algorithm provides smaller errors in the forecast and better staffing levels in call centers. The results can be used for a more flexible workforce management in call centers.

24. H. Shen, J. Z. Huang. *Analysis of Call Centre Arrival Data Using Singular Value Decomposition*, Applied Stochastic Models in Business and Industry, May 2005.
http://onlinelibrary.wiley.com/doi/10.1002/asmb.598/pdf

Abstract: We consider the general problem of analysing and modelling call centre arrival data. A method is described for analysing such data using singular value decomposition (SVD). We illustrate that the outcome from the SVD can be used for data visualization, detection of anomalies (outliers), and extraction of significant features from noisy data. The SVD can also be employed as a data reduction tool. Its application usually results in a parsimonious representation of the original data without losing much information. We describe how one can use the reduced data for some further, more formal statistical analysis. For example, a short term forecasting model for call volumes is developed, which is multiplicative with a time series component that depends on day of the week. We report empirical results from applying the proposed method to some real data collected at a call centre of a large-scale U.S. financial organization. Some issues about forecasting call volumes are also discussed.

25. E. Chassioti, D. J. Worthington. *A New Model for Call Centre Queue Management*, The Journal of the Operational Research Society, Vol. 55, No. 12 (Dec., 2004), pp. 1352–1357. http://www.jstor.org/stable/pdfplus/4101854.pdf?acceptTC=true

Abstract: A new model for call centre queue management is described. It incorporates important features of call centre queues and is shown to produce results that are very different from those produced by the more usual models. The analytic approach is easy to apply, and is used to offer some interesting insights for call center queue management.

26. Yen Chu Cheng. *Analysis of Call Center Data*, April 2004. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1002&context=wharton_research_scholars

Abstract: A call center is a place where a group of agents service customers remotely via the telephone. Queueing theory is used extensively in the study of call centers. One of the most important statistical models used to predict call arrivals in queueing theory is the Poisson process. This paper summarizes the results of the application of this model in the study of call arrivals to a particular bank. Specifically, the call arrivals to the VRU (Voice Response Unit) and to the service queue are examined. These calls are put through different techniques that test whether the actual data conform to the hypothetical model for a Poisson process. An explanation of each of these tests and the results are described in this paper.

27. L. D. Broun, L. H. Zhao. *A Test for the Poisson Distribution,* Sankhya: The Indian Journal of Statistics, 2002, Vol. 64, Series A, Pt. 3, pp. 611–625. http://sankhya.isical.ac.in/search/64a3/64a3037.pdf

Abstract: We consider the problem of testing whether a sample of observations comes from a single Poisson distribution. Of particular interest is the alternative that the observations come from Poisson distributions with different parameters. Such a situation would correspond to the frequently discussed situation of overdispersion. We propose a new test for this problem that is based on Anscombe's variance stabilizing transformation. There are a number of tests commonly proposed, and we compare the performance of these tests under the null hypothesis with that of our new test. We find that the performance of our test is competitive with the two best of these. The asymptotic distribution of the new test is derived and discussed. Use of these tests is illustrated through two examples of analysis of call-arrival times from a telephone call center. The example facilitates careful discussion of the performance of the tests for small parameter values and moderately large sample sizes.

28. Søren Asmussen, Olle Nerman, Marita Olsson. *Fitting Phase-type Distributions via the EM Algorithm*, Scandinavian Journal of Statistics, Vol. 23, No. 4 (Dec, 1996), pp. 419–441. (This research uses data that was later absorbed into SEE repositories.) http://www.jstor.org/stable/4616418

Abstract: Estimation from sample data and density approximation with phase-type distributions are considered. Maximum likelihood estimation via the EM algorithm is discussed and performed for some data sets. An extended EM algorithm is used to minimize the information divergence (maximize the relative entropy) in the density approximation case. Fits to Weibull, log normal, and Erlang distributions are used as illustrations of the latter.