Performance Evaluation 78 (2014) 1–6

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

Short communication

MDP based optimal pricing for a cloud computing queueing model☆

Rami Atar, Israel Cidon, Mark Shifrin *

Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 4 January 2014
Received in revised form 5 May 2014
Accepted 5 May 2014
Available online 27 May 2014

Keywords:
Optimal pricing
Price-based revenue management
MDP
Threshold policy
Cloud computing

ABSTRACT

We study the optimization of dynamic pricing in a queueing model with a finite buffer, where arrival rates depend on advertised price levels. We apply our study to a pricing policy in a cloud computing service provider setup.

The main result of this paper is the multi-threshold structure of the optimal policy.

1. Introduction

The model under study consists of a finite buffer and a single processing unit. Tasks of multiple types arrive into the system and are admitted unless the buffer is full. Tasks that are not admitted are lost. Price levels charged per task are advertised for the various types of tasks. It is assumed that the arrival rates of tasks to the system depend on these advertised prices. The decision maker (DM) dynamically controls the price levels and attempts to maximize an expected discounted total reward. This type of problem is often referred to in the literature as *price-based revenue management* (see [1]). Our main result is the identification of an optimal policy that has a threshold structure. While the formulation of this problem via Markov decision processes (MDP) is standard, the optimality of a threshold policy for this model does not seem to follow from existing results. Our motivation for studying this model stems from possible application areas such as computer network managing systems [2,3] and especially cloud computing services at both consumer and enterprise markets [4,5]. Previous MDP formulations for cloud computing applications can be found in [6,7]. These works optimize the decision making of the cloud computing customer.

The area of optimal control of queueing models via MDP has been studied extensively. Notable examples are [8–11], and the monographs [12,13]. Papers such as [9,10] provide an explicit characterization of optimal control as a threshold policy. The way to establish this property in these papers and many others is through the convexity of the value function. In our paper this property does not follow from convexity but from a different condition involving second order differences (see Section 3).

Among works that address queueing models with finite buffers let us mention [14] that addresses optimal task admission in the scheduler-to-router system, [15] that calculates individual and social optimum in a problem of admission into a finite

☆ Research supported in part by the ISF (Grant 1315/12), the US–Israel BSF (Grant 2008466), and the Technion fund for promotion of research.
* Corresponding author. Tel.: +972 546406433.
E-mail address: shifrin@tx.technion.ac.il (M. Shifrin).

http://dx.doi.org/10.1016/j.peva.2014.05.001
0166-5316/

buffer, and [16] that addresses the optimization of capital expense of the buffer size. Some recent works on optimal pricing problems based on queueing models include [17,18].

Our main motivation for the model under consideration comes from pricing in cloud computing. The model described here conforms with the *pay-as-you-go* pricing method; see [5] for this term as well as a review of pricing models in this context.

In Sections 2 and 3 we present the model, and respectively, the main result and its proof. In Section 4 we provide more motivation for the cloud computing model and compute a possible pricing policy based on the actual price list of Amazon Web Services (AWS).

2. Model and main result

The main result is stated in this section for a single task type; a generalization to multiple types will be addressed in Section 3.2. The model considered has K price levels

$$C_1 < C_2 < \cdots < C_K,$$

that are subject to the dynamic selection by a DM. These decisions influence the arrival rates. Namely, advertising the price C_k gives rise to arrival rate λ_k, where

$$\lambda_1 > \lambda_2 > \cdots > \lambda_K.$$

It is further assumed that the buffer can contain no more than B tasks at any given time.

Let $E_k, k \in \{1, \ldots, K\}$ be independent Poisson processes of intensities λ_k, respectively. Denote by A the counting processes for arrivals. Denote by $A_k(t)$ the number of tasks priced at level k that has arrived up to time t. It is assumed that

$$A_k = \int_0^t U_k(s) dE_k(s),$$

where $\{U_k, 1 \le k \le K\}$ is an S-valued process, where

$$\mathrm{S} = \left\{ u \in \{0, 1\}^K : \sum u_k \le 1 \right\}.$$

Advertising a price C_k at time t corresponds to selecting $U_j(t) = 1$ for $j = k$ and $U_j(t) = 0$ for all other j. The option $U_k(t) = 0$ for all k is also possible, and corresponds to rejection, which by assumption can occur only when the buffer is full. We regard U as the control process.

The total number of tasks present in the buffer at time t is given by

$$X(t) = x + A(t) - D(t), \tag{1}$$

where $x \in \{0, \ldots, B\}$ denotes an initial condition, and D is the departure process, counting the number of completed tasks of all types. The service time distribution is assumed to be exponential of rate $\mu > 0$. The control process U is regarded admissible if it is adapted to the filtration generated by $(\{E_k\}, X)$, and is such that the buffer limit $X(t) \le B$ is kept at all times. The class of admissible control processes is denoted by $\mathcal{U}$.

We consider an expected discounted reward given by

$$\begin{aligned} J(x, U) &= E\Big[\int_0^\infty e^{-\gamma s} \sum_{k=1}^K C_k dA_k(s)\Big] \\ &= E\Big[\int_0^\infty e^{-\gamma s} \sum_{k=1}^K C_k U_k(s) dE_k(s)\Big], \end{aligned} \tag{2}$$

for a fixed $\gamma > 0$. The value function is defined as

$$V(x) = \sup_{U \in \mathcal{U}} J(x, U), \quad x \in \{0, 1, \ldots, B\}. \tag{3}$$

Theorem 2.1. *There exist constants*

$$0 = b_0 \le b_1 \le \cdots \le b_K = B + 1,$$

such that the following policy is optimal: announce price C_i at time t if and only if $b_{i-1} \le X(t) < b_i$.

3. Proof and extensions

3.1. Proof

Denote $\delta_i = (\mu + \lambda_i + \gamma)^{-1}$ and note that $\delta_1 < \delta_2 < \cdots < \delta_K$. Also observe $\delta_1\lambda_1 > \delta_2\lambda_2 > \cdots > \delta_K\lambda_K$. For $i < j$, denote $\beta_{ij} = \lambda_i\delta_i - \lambda_j\delta_j$ and $\alpha_{ij} = \mu\delta_j - \mu\delta_i$. Denote $\bar{\delta} = (\gamma + \mu)^{-1}$. Observe that α_{ij} and β_{ij} $(i < j)$ are all positive. Next, for $i < j < k$, we have

$$\beta_{ij} + \beta_{jk} = \beta_{ik}, \qquad \alpha_{ij} + \alpha_{jk} = \alpha_{ik}. \tag{4}$$

Moreover,

$$\alpha_{ij} = \frac{\mu\left(\lambda_i - \lambda_j\right)}{(\gamma+\mu+\lambda_i)\left(\gamma+\mu+\lambda_j\right)} \tag{5}$$

$$\beta_{ij} = \frac{\lambda_i\gamma + \lambda_i\mu - \lambda_j\gamma - \lambda_j\mu}{(\gamma+\mu+\lambda_i)\left(\gamma+\mu+\lambda_j\right)}. \tag{6}$$

We will use the notation

$$\tilde{V}_{ij}(x) = \beta_{ij}V(x+1) - \alpha_{ij}V(x-1),$$

and

$$W_i(x) = \delta_i\mu V(x-1) + \delta_i\lambda_i(V(x+1)+C_i).$$

The value function uniquely solves the Bellman equation (see e.g. [13, Chapter 8])

$$V(x) = \max_j\{\delta_j\mu V(x-1) + \delta_j\lambda_j(V(x+1)+C_j)\}, \quad x \in \{1,2,\dots,B-1\},\ j \in \{1,\dots,K\}, \tag{7}$$

with the boundary conditions

$$V(0) = \max_j\{\delta_j\mu V(0) + \delta_j\lambda_j(V(1)+C_j)\}, \quad j \in \{1,\dots,K\}, \tag{8}$$

$$V(B) = \bar{\delta}\mu V(B-1). \tag{9}$$

For a function $V : \{0,1,\dots,B\} \to \mathbb{R}$, consider the property

$$\tilde{V}_{ij}(x) := \beta_{ij}V(x+1) - \alpha_{ij}V(x-1) \quad \text{is nonincreasing in } x \in \{1,\dots,B-1\}, \text{for each } i,j,\ 1 \le i < j \le K. \tag{10}$$

We will argue that V has this property. To this end, consider the operator T, acting in the space of functions from $\{0,1,\dots,B\}$ to $\mathbb{R}$, defined as

$$TU(x) = \max_j\{[\delta_j\mu U(x-1) + \delta_j\lambda_j(U(x+1)+C_j)]\} \quad x \in \{1,\dots,B-1\}, \tag{11}$$

$$TU(B) = \bar{\delta}\mu U(B-1),$$

$$TU(0) = \max_j\{[\delta_j\mu U(0) + \delta_j\lambda_j(U(1)+C_j)]\},$$

for $U : \{0,1,\dots,B\} \to \mathbb{R}$. Then the Bellman equation reads $TV = V$. Let S be the set of functions $U : \{0,1,\dots,B\} \to \mathbb{R}$ that are non-increasing and possess the property (10). The following lemma asserts that T preserves S, and moreover, with $\|U\| := \max_x |U(x)|$, acts on it as a strict contraction.

Lemma 3.1. *One has $TS \subset S$. Moreover, there exists a constant $a \in (0,1)$ such that*

$$\|TU - TW\| \le a\|U - W\| \quad \text{for every } U, W \in S.$$

Proof. To prove the first assertion, let $U \in S$ be given. For $x \in \{2,\dots,B-1\}$, $TU(x) - TU(x-1) \le 0$ by (11), using the fact that the maximum of two nonincreasing functions is nonincreasing. A verification for $x = 1$ and $x = B$ gives $TU(x) - TU(x-1) \le 0$ as well, and the nonincreasing property of TU follows.

To prove the contraction property, let $U, W \in S$. Consider first $x \in \{1,2,\dots,B-1\}$. We have

$$TU(x) - TW(x) = \max_j\{[\delta_j\mu U(x-1) + \delta_j\lambda_j(U(x+1)+C_j)]\} - \max_i\{[\delta_i\mu W(x-1) + \delta_i\lambda_i(W(x+1)+C_i)]\}. \tag{12}$$

Using

$$|\max(a_1,a_2,\dots,a_K) - \max(b_1,b_2,\dots,b_K)| \le \max(|a_1-b_1|,\dots,|a_K-b_K|)$$

and denoting $\Delta = U - W$,

$$|TU(x) - TW(x)| \le \max_k\{|\delta_k\mu\Delta(x-1) + \delta_k\lambda_k\Delta(x+1)|\}.$$

Since $\delta_k\mu + \delta_k\lambda_k < 1$, this gives

$$|TU(x) - TW(x)| \le a\|\Delta\| = a\|U - W\|,$$

where $a < 1$. A similar calculation for $x = 0$ and $x = B$ gives an analogous inequality, and we conclude that $\|TU - TW\| \le a\|U - W\|$.

Proving the property (10) of TV amounts to showing that

$$\beta_{ij}TV(x+1) - \alpha_{ij}TV(x-1) \quad \text{is nonincreasing in } x, \text{ for each } i,j,\ 1 \le i < j \le K. \tag{13}$$

We show that

$$\beta_{ij}W_k(x+1)-\alpha_{ij}W_l(x-1) \quad \text{is nonincreasing for } 1\le i<j\le K,\ 1\le l\le k\le K. \tag{14}$$

Due to the nonincreasing property of $W_k(x)-W_l(x)$, the index k of $W_k(x+1)$ is always greater than or equal to the index l of $W_l(x-1)$ in (14). Thus there is no need to consider all other cases. We use this observation for the proof of the property. Expand (14) as follows:

$$\beta_{ij}\delta_k\lambda_kV(x+2)+\beta_{ij}\delta_k\mu V(x)-\alpha_{ij}\delta_l\lambda_lV(x)-\alpha_{ij}\delta_l\mu V(x-2)+C$$

where C is a constant. Omitting C rewrite this expression as

$$\beta_{ij}\delta_l\lambda_lV(x+2)+\beta_{ij}\delta_l\mu V(x)+\beta_{ij}(\delta_k\lambda_k-\delta_l\lambda_l)V(x+2)+\beta_{ij}(\delta_k\mu-\delta_l\mu)V(x)-\alpha_{ij}\delta_l\lambda_lV(x)-\alpha_{ij}\delta_l\mu V(x-2)$$
$$=\lambda_l\delta_l(\beta_{ij}V(x+2)-\alpha_{ij}V_x)+\delta_l\mu\tilde{V}_{ij}(x-1)-\beta_{ij}\beta_{lk}V(x+2)+\alpha_{lk}\beta_{ij}V(x).$$

Note that by assumption, $\delta_l\mu\tilde{V}_{ij}(x-1)$ is nonincreasing. We show that the remaining part of the above display, namely

$$(\beta_{ij}\delta_l\lambda_l-\beta_{ij}\beta_{kl})V(x+2)+(\beta_{ij}\alpha_{lk}-\alpha_{kl}\delta_l\lambda_l)V(x),$$

is nondecreasing as well. Using (5) and (6) we obtain

$$\frac{\left(-\lambda_i\mu+\lambda_j\mu\right)\lambda_kV(x)}{(\gamma+\mu+\lambda_k)(\gamma+\mu+\lambda_i)\left(\gamma+\mu+\lambda_j\right)}+\frac{\left(\lambda_i\gamma+\lambda_i\mu-\lambda_j\gamma-\lambda_i\mu\right)\lambda_kV(x+2)}{(\gamma+\mu+\lambda_k)(\gamma+\mu+\lambda_i)\left(\gamma+\mu+\lambda_j\right)},$$

that further reduces to $V(x+2)\beta_{ij}\lambda_k\delta_k-V(x)\alpha_{ij}\lambda_k\delta_k=\lambda_k\delta_k\tilde{V}_{ij}(x+1)$. We thus have

$$\beta_{ij}TV(x+1)-\alpha_{ij}TV(x-1)=\lambda_k\delta_k\tilde{V}_{ij}(x+1)+\delta_l\mu\tilde{V}_{ij}(x-1),$$

that is the sum of two nonincreasing functions and as such is nonincreasing as well. This completes the proof of the lemma.

Proof of Theorem 2.1. We use the contraction mapping principle (see e.g. [19, Theorem V.18]). The set S, equipped with the metric $\rho(U,W)=\|U-W\|$ is a complete metric space. The map $T:S\to S$ is a strict contraction, as shown in the above lemma. As a result, T has a unique fixed point. That is, there exists a unique $U\in S$ for which $TU=U$. Recall that V is the unique solution to the same equation in the space of *all* functions from $\{0,1,\dots,B\}$ to $\mathbb{R}$. As a result, $V=U$. This shows $V\in S$, namely, that V is nonincreasing and possesses the property defined in (10). Thus, we deduce that $V\in S$.

Finally, the optimal action at state x can be read from the Bellman equation (7). The action depends on finding index i for which the inequality

$$\delta_i\mu V(x-1)+\delta_i\lambda_i(V(x+1)+C_i)>\delta_j\mu V(x-1)+\delta_j\lambda_j(V(x+1)+C_j)$$

holds for all $j\neq i$. This can be written as

$$\beta_{ij}V(x+1)-\alpha_{ij}V(x-1)>C:=-C_i\delta_i\lambda_i+C_j\delta_j\lambda_j,$$

namely $\tilde{V}_{ij}(x)>C$. The monotonicity property of $\tilde{V}_{ij}$ thus gives the threshold property and the result follows.

3.2. System with multiple task types

The result can be extended to cover multiple task types. Consider a system with I task types, where each type, i, arrives at the system with rate $\lambda_{i,k}$ whenever price $c_{i,k}$ is advertised. It is assumed that $\lambda_{i,k}$ (resp., $c_{i,k}$) are non-increasing (resp., non-decreasing) in k, for each i. Here, k varies over $\{1,\dots,K\}$. It is also assumed that tasks of all types are served at the same rate μ.

Write (j) for a generic multi-index $(j_1,\dots,j_I)$, where each j_i takes values in $\{1,\dots,K\}$. The Bellman equation takes the following form:

$$V(x)=\max_{(j)}\{[\delta_{(j)}\mu V(x-1)+\delta_{(j)}\Lambda_{(j)}(V(x+1)+C_{(j)})]\},$$

where

$$\Lambda_{(j)}=\sum_{i=1}^{I}\lambda_{i,j_i},\qquad C_{(j)}=\frac{1}{\Lambda_{(j)}}\sum_{i=1}^{I}c_{i,j_i}\lambda_{i,j_i},$$

and $\delta_{(j)}=(\Lambda_{(j)}+\mu+\gamma)^{-1}$. Label the multi-indices (j) by $n\in\{1,\dots,N\}$, where $N=K^I$, in such a way that $\delta_1\le\delta_2\le\cdots\le\delta_N$ (equivalently, $\delta_1\lambda_1\ge\delta_2\lambda_2\ge\cdots\ge\delta_N\lambda_N$.) An extension of Theorem 2.1 to the present setting then states the following:

There are constants $0=b_0\le b_1\le\cdots\le b_N=B+1$, *such that it is optimal to announce at time* t *the price combination* $C_{(j)}=(c_{j_1},\dots,c_{j_I})$ *if and only if* $b_{n-1}\le X(t)<b_n$.

Here n is the label assigned to the multi-index (j), and X represents the total number of tasks present in the system.

Table 1
Cloud with dynamically advertised prices—threshold examples.

B	λ_i	C_i	b_i
20	3 2 1	2.0 2.1 2.2	3 10 20
30	1.5 1.4 1	2.0 2.1 2.2	2 35 50
100	3.5 3.0 2.0 1.5	2.5 2.6 2.75 2.8	48 55 57 60
200	0.5 0.4 0.2 0.1	1.6 2.0 3.75 7.5	193 199 199 200

4. Example: optimizing cloud provider revenue

We apply the framework presented above to a setting where a cloud service provider optimizes its revenue. If a certain amount of cloud resources is not leased by subscribed customers, it makes sense for the cloud provider to offer these unused resources for opportunistic on-demand usage (cloudbursting). In this setup dynamic pricing can significantly increase the revenue. As the demand for cloudbursting decreases, the cloud provider can apply a price reduction to compete for a larger share of the customers. In times of high cloudbursting demands the service provider may prefer to select the most profitable task types. The finite buffer in our model reflects a quality of service requirement. The cloud must serve accepted tasks within certain average delay.

A numerical result of pricing policies based on the MDP approach is presented in Table 1. The prices were taken from Amazon EC2 [20] *spot instances* price list. The pricing here is for leasing a single computational resource—the spot instance per hour, to which the data transfer and storage costs are added up. The spot instances are charged at spot prices. The prices are regulated and fluctuate periodically depending on the availability and demand of resources. We derive the buffer limit for a scenario where the delay constraint is 720 h (1 month), and we vary the average task processing time. For example, in the first row the average task processing rate is 36 h (thus $\mu = 1/36$ tasks/hour), and so B was set to $720/36 = 20$ tasks. We normalized the task processing rate and the arrival rates such that the task processing rate was always set 1. Thus, in the first line the rates were multiplied by 36.

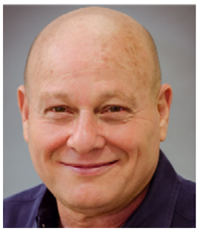
The structure of the table is as follows. The first three columns are the possible proposed quantities, measured in tasks, normalized tasks per hour, and price per task. Finally, the last column gives the pricing thresholds.

Note that in the example in the last row, the third price level is never optimal.

References

[1] K.T. Talluri, G.J. Van Ryzin, Revenue Management, Sprintgter Science and Business Media, NY, 2004.
[2] I. Mitrani, Management of server farms for performance and profit, Comput. J. 53 (7) (2010) 1038–1044.
[3] S. Ren, J. Park, M. van der Schaar, User subscription dynamics and revenue maximization in communications markets, in: INFOCOM, 2011 Proceedings IEEE, IEEE, 2011, pp. 2696–2704.
[4] M.A. Salehi, R. Buyya, Adapting market-oriented scheduling policies for cloud computing, in: Algorithms and Architectures for Parallel Processing, Springer, 2010, pp. 351–362.
[5] H. Wang, Q. Jing, R. Chen, B. He, Z. Qian, L. Zhou, Distributed systems meet economics: pricing in the cloud, in: Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing, USENIX Association, 2010.
[6] M. Shifrin, Admission and Scheduling Control in Cloud Computing-Markov Decision Processes and Diffusion Approximations (Ph.D. thesis), Technion, 2013.
[7] M. Shifrin, R. Atar, I. Cidon, Optimal scheduling in the hybrid-cloud, in: 2013 IFIP/IEEE International Symposium on Integrated Network Management, (IM 2013), IEEE, 2013, pp. 51–59.
[8] E.B. Çil, E.L. Örmeci, F. Karaesmen, Effects of system parameters on the optimal policy structure in a class of queueing control problems, Queueing Syst. 61 (4) (2009) 273–304.
[9] B. Hajek, Optimal control of two interacting service stations, IEEE Trans. Automat. Control 29 (6) (1984) 491–499.
[10] W.E. Helm, K.-H. Waldmann, Optimal control of arrivals to multiserver queues in a random environment, J. Appl. Probab. (1984) 602–615.
[11] R.R. Weber, S. Stidham Jr., Optimal control of service rates in networks of queues, Adv. Appl. Probab. (1987) 202–218.
[12] G. Koole, Monotonicity in Markov Reward and Decision Chains: Theory and Applications, Vol. 1, Now Publishers Inc., 2007.
[13] J. Walrand, An Introduction to Queueing Networks, Vol. 165, Prentice Hall, Englewood Cliffs, NJ, 1988.
[14] K.Y. Lin, S.M. Ross, Admission control with incomplete information of a queueing system, Oper. Res. 51 (4) (2003) 645–654.
[15] U. Sumita, Y. Masuda, S. Yamakawa, Optimal internal pricing and capacity planning for service facility with finite buffer, European J. Oper. Res. 128 (1) (2001) 192–205.
[16] B.L. Miller, A. Buckman, Cost allocation and opportunity costs, Manag. Sci. 33 (5) (1987) 626–639.
[17] R. Atar, M.I. Reiman, Asymptotically optimal dynamic pricing for network revenue management, Stoch. Syst. 2 (2012) 232–276.
[18] U. Yildirim, J.J. Hasenbein, Admission control and pricing in a queue with batch arrivals, Oper. Res. Lett. 38 (5) (2010) 427–431.
[19] M.C. Reed, B. Simon, Methods of Modern Mathematical Physics: Functional Analysis. 1, Vol. 1, Access Online via Elsevier, 1980.
[20] Amazon EC2 pricing. http://aws.amazon.com/ec2/pricing/. Amazon Web Services.

Rami Atar is a professor at the Electrical Engineering Department, Technion. He received the D.Sc. degree in Electrical Engineering from the Technion in 1997. After postdoc positions at the Division of Applied Mathematics at Brown and the Fields Institute, he joined the Faculty of Electrical Engineering, Technion, in 1999. He had visiting appointments with the University of Washington and the University of North Carolina. He is a fellow of the IMS. Atar's research interests are in stochastic processes, including asymptotic analysis of queueing models in diffusion and large deviation regimes, PDE techniques in stochastic control and differential games, filtering and estimation.

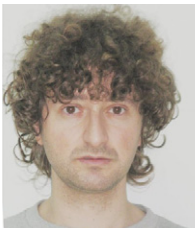


Israel Cidon is a chaired professor and a former dean of Electrical Engineering at the Technion—Israel Institute of Technology.

In 2012, he co-founded Sookasa that provides a layer of compliance and security for popular cloud services. In 2000, he co-founded Actona Technologies that pioneered the fields of Wide Area File Systems (WAFS) and WAN optimization. Actona was acquired by Cisco in 2004 and is the source of Cisco, Wide Area Application Services (WAAS). In 1998 he co-founded Viola Networks, the provider of Netally, a distributed active network testing suite that was acquired by Fluke Networks in 2008. In 1981 he co-founded Micronet, an early vendor of mobile computers (Israeli IPO 2006).

In 1985–1994 he was a research member and the manager of the Network Architecture group at IBM Research in NY, leading the first implementations of triple-play packet network and metropolitan optical resilient packet rings. He received the IBM outstanding innovation award in 1989 and 1993. In 1994–1995 he led high-speed networking at Sun Labs, Mountain View., working on open network control architectures.

His academic research covers aspects of communication networks including mobility, wireless, Internet and Network on Chip (NoC). He is a co-author of over 170 refereed papers and 27 US patents.



Mark Shifrin, received the PhD degree from Faculty of Electrical Engineering, Technion in 2014.

His research interests include Markov decision processes, queuing theory, analysis of queueing models in diffusion regime, and stochastic control and their applications in computer networks.

He is a recipient of Kreitman Post-Doctoral Scholarship from Ben Gurion University.