UDC 519.872.5, 519.216.24

Fluid Limit for Closed Queueing Network with Several Multi-Servers

S. V. Anulova*

* *V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences*
65 Profsoyuznaya street, Moscow, 117997, Russia

Abstract. A closed network consists of several multi-servers with n customers. Service requirements of customers at a multi-server have a common cdf. State parameters of the network: for each multi-server empirical measure of the age of customers being serviced and for the queues the numbers of customers in them, all multiplied by n^{-1}.

Our objective: asymptotics of dynamics as $n \to \infty$. The asymptotics of dynamics of a single multi-server and its queue with an arrival process as the number of servers $n \to \infty$ is currently studied by famous scientists K. Ramanan, W. Whitt et al. Presently there are no universal results for general distributions of service requirements—the results are either for continuous or for discrete time ones; the same for the arrival process. We establish the asymptotics for a network in discrete time, find its equilibrium and prove convergence as $t \to \infty$.

Motivation for studying such models: they represent call/contact centers and help to construct them effectively.

Keywords: call/contact centers, queueing network, fluid limit approximation.

1. Introduction

1.1. Review of Investigated Contact Centers Models

In the last 15 years an extensive research in mathematical models for telephone call centers has been carried out, cf. References [2-6, 8–17] of [1]. For the state process were found fluid limits as the number of servers tends to infinity.

An important particular question is the convergence of the fluid limit to a stable state as time tends to infinity. For a discrete time model W. Whitt has found equilibrium states (a multitude) of the fluid model and proved the time convergence in a special case—for a primitive arrival process and for initial condition with empty multi-server and queue, [3, Section 7].

1.2. A New Model for Contact Centers and Its Fluid Limit with Equilibrium Behavior

We have suggested in [1] a more suitable model for contact centers. The number of customers is fixed. Customers may be situated in two states: normal and failure. There is a multi-server which repairs customers in the failure state. The repair time/the time duration of a normal state is a random variable, independent and identically distributed for all customers. For a large number of customers and a suitable number of servers we have calculated approximately the dynamics of the normalized state of the system—its fluid limit. And in [2] we explored the convergence of the fluid limit as time tends to infinity and found its steady-state (or equilibrium). Now we establish the described properties for a generalized model: there are several repair multi-servers.

2. Closed Multi-Server Network with n Customers and Its Fluid Limit Equilibrium

2.1. Network Description

Consider a closed network consisting of n customers and N multi-servers. Multi-server 1 (further denoted MS1) consists of n servers (for the customers in the normal state), the time they service a customer has distribution G^1. Multi-server i (further denoted MSi) consists of $s_n^i n$ servers with a number $s_n^i \in (0,1)$ (for the customers in the failure state type i), the time they service a customer has distribution G^i, $i = 2, \ldots, N$. The distributions G^i, $i = 1, \ldots, N$, are discrete: they are concentrated on $\{1, 2, \ldots\}$. Customers move from MSi to MS1, and from MS1 to MSi with probability p_i, $i = 1, \ldots, N$. Service times are independent for all servers and all customers. We will investigate the behavior of the network as $n \to \infty$, in discrete time $t = 0, 1, 2, \ldots$

In MS1 no queue may arise—if all n its servers are occupied then all customers are in MS1, therefore no new customer can arrive.

Denote the number of customers at a moment $t = 0, 1, \ldots$ in MSi by $B_n^i(t)$ and in its queue by $Q_n^i(t)$, $i = 1, \ldots, N$, and the number of customers in the whole queue $Q_n(t) = \sum_{i=1}^n Q_n^i(t) = n - \sum_{i=1}^n B_n^i(t)$.

$$B_n^i(t) = \sum_{k=0}^{\infty} b_n^i(t,k) \text{ and } (s_n^i - B_n^i(t)/n) Q_n^i(t) = 0 \text{ for all } t \text{ and } n,$$

with $b_n^i(t,k)$ being the number of customers in the multi-server i at the moment t who have spent there time k, $i = 1, \ldots, N$.

2.2. Fluid Limit Dynamics

Notations

Denote for $i = 1, \ldots, N$:

— $G^{i;c}(k) := 1 - G^i(k)$ and $g^i(k) := G^i(k) - G^i(k-1), k = 1, 2, \ldots$

— E^i the expectation of the time the server in MSi services a customer:

$$E^i := \sum_{k=1}^{\infty} k g^i(k) = 1 + \sum_{k=1}^{\infty} G^i(k).$$

— $\sigma_n^i(t)$ the number of service completions in MSi at time moment $t = 1, 2, \dots$

Fluid limit dynamics

Under certain conditions, specifically $\lim_{n\to\infty} s_n^i = s^i \in (0,1)$, $i = 2,\dots,N$, the fluid limit exists and its dynamics is described below (the proof for $N = 2$ is given in [1, Theorem 1]). As $n \to \infty$, for $i = 1,\dots,N$ and t, $k \geq 0$,

$$\frac{b_n^i(t,k)}{n} \Rightarrow b^i(t,k),\ \frac{\sigma_n^i(t)}{n} \Rightarrow \sigma^i(t),\ \frac{B_n^i(t)}{n} \Rightarrow B^i(t) \equiv \sum_{k=0}^{\infty} b^i(t,k) \geq 0,$$

$$\frac{Q_n^i(t)}{n} \Rightarrow Q^i(t),\ Q(t) \equiv \sum_{i=1}^{N} Q^i \geq 0,\ \sum_{i=1}^{N} B^i(t) + Q(t) = 1,$$

and for $i = 2,\dots,N$ $\quad B^i(t) \leq s^i$ and $(s^i - B^i(t))Q^i(t) = 0$. The evolution of the vector $(b^i, \sigma^i,\ i = 1,\dots,N)(t)$, $t = 0, 1\dots$, proceeds with steps of t:

$$b^i(t,k) = b^i(t-1,k-1)\frac{G^{i;c}(k)}{G^{i;c}(k-1)},\ k = 1,2,\dots,\ i = 1,\dots,N,$$

$$b^i(t,0) = \min\{s^i - B^i(t-1) + \sigma^i(t), Q^i(t-1) + p_i\sigma^1(t)\},\ i = 2,\dots,N,$$

$$b^1(t,0) = \sum_{i=2}^{N} \sigma^i(t),\ \sigma^i(t) = \sum_{k=1}^{\infty} b^i(t-1,k-1)\frac{g^i(k)}{G^{i;c}(k-1)}.$$

2.3. Fluid Limit Equilibrium

Consider the discrete time fluid limit for the closed network model dynamics described in subsecton 2.2.

Definition 1. *A point in the state space of deterministic fluid processes is called "equilibrium" if fluid processes after reaching this point remain in it. Equilibrium points are described/characterized by sets* $(b^{*i} = \{b^{*i}(k), k = 0,1,\dots\},\ Q^{*i},\ i = 1,\dots,N)$ *satisfying*

$$\sum_{k=0}^{\infty} b^{*i}(k) \leq s^i,\ i = 2,\dots,N, \quad \sum_{i=1}^{N}\left(\sum_{k=0}^{\infty} b^{*i}(k) + Q^{*i}\right) = 1.$$

Denote $B^{*i} = \sum_{k=0}^{\infty} b^{*i}(k)$, $Q^* = \sum_{i=1}^{N} Q^{*i}$, $i = 1,\dots,N$.

Theorem 1. *For the deterministic fluid processes there exists a nearly single equilibrium point. The characteristics* b^{*i}, $B^{*i}, i \in \{1,2,\dots,N\}, Q^*$ *of this equilibrium point have the form:*

1. $b^{*i}(0) = p_i b^{*1}(0)$, $i \in \{2, \dots, N\}$.
2. $b^{*i}(k) = b^{*i}(0) G^{i;c}(k)$, $k = 1, 2, \dots$, $i \in \{1, 2, \dots, N\}$.
3. $B^{*i} = b^{*i}(0) E^i$, $i \in \{1, 2, \dots, N\}$.
4. $Q^* = \sum_{i=2}^{N} Q^{*i} = 1 - \sum_{i=1}^{N} B^{*i}$.

Denote $L := \{i \in \{2, \dots, N\} \mid \quad \dfrac{p_i E^i}{\sum_{i=1}^{N} p_i E^i} > s^i\}$. *Then*

$$b^{*1}(0) = \frac{1}{\sum_{i=1}^{N} p_i E^i} \; \textit{if } L = \emptyset, \quad \textit{or } \min_{i \in L} \frac{s^i}{p_i E^i} \; \textit{if } L \neq \emptyset,$$

and $\sum\limits_{i \notin L} Q^{*i} = 0, \quad \sum\limits_{i \in L} Q^{*i} = Q^*$ *with any selection of* $\{Q^{*i}, \, i \in L\}$.

Corollary 1. *If in the practice it is desirable to organize the call/contact center with quick services for customers, that is, without queues, then each multi-server must be large enough:* $s^i \geq \dfrac{p_i E^i}{\sum_{i=1}^{N} p_i E^i}$, $i = 2, \dots, N$.

2.4. Fluid Limit Convergence to Equilibrium As $t \to \infty$

No strong result for universal convergence has been presented by W. Whitt in [3, Section 7], only starting from an empty multi-server and an empty queue. We shall transfer this simple theorem to our closed network model.

If MS1 is empty, MSi is filled with equilibrium parameters, where $i \in \{2, \dots, N\}$, $\sum\limits_{i \notin L} Q^i = 0$, then MSi, $i \in \{2, \dots, N\}$, remain in this state, the queues decrease and MS1 adds with time steps customers of the next age with equilibrium parameters, and the state of MS1 converges monotonically to the unique equilibrium state: for $t = 1, 2, \dots$

$$b^1(t, k) = \begin{cases} b^{*1}(0) G^{1;c}(k), & 0 \leq k < t, \\ 0, & k \geq t. \end{cases}$$

Theorem 2. *Suppose the fluid limit satisfies at time* $t = 0$ *the following conditiions:* $B^1(0) = 0$ *and* $b^i(0, \cdot) = b^{*i}$, $i \in \{2, \dots, N\}$. *Then the fluid limit converges to the equilibrium point as* $t \to \infty$. *Namely:*

— *the state of MSi remains equilibrium:*

$$b^i(t, \cdot) = b^{*i}, \; t = 0, 1, 2, \dots, , i \in \{2, \dots, N\};$$

— *the state of MS1 grows occupying its equilibrium state—with each time step adds the next age equilibrium parameter:*

$$b^1(0, \cdot) \equiv 0 \textit{ and for } t = 1, 2, \dots \; b^1(t, \cdot) = b^{*1}(\cdot) I_t \quad \textit{with } I_t = I_{\{0,1,\dots,t-1\}};$$

— *the queue decreases—with each time step loses the amount of the previous age MS1 equilibrium parameter:*

$$Q(0) = 1 - \sum_{2}^{N} B^{*i}, \quad Q(t) = Q(t-1) - b^{*1}(t-1)$$

$$= 1 - \sum_{2}^{N} B^{*i} - \sum b^{*1}(\cdot) I_t = 1 - \sum_{2}^{N} B^{*i} - \sum_{l=0}^{t-1} b^{*1}(l),$$

and $Q^i(t) = Q^i(t-1) - p_i b^{*1}(t-1)$, $t = 1, 2, \ldots$, $i \in \{2, \ldots, N\}$.

Acknowledgments

This work was partially supported by RFBR grants No. 16-08-01285 A "Control of stochastic, deterministic, and quantum systems in phases of quick movement." and No. 17-01-00633 A "Problems of stability and control in stochastic models".

References

1. *Anulova S.* Approximate description of dynamics of a closed queueing network including multi-servers // 18th International Conference "Distributed Computer and Communication Networks", Moscow, Russia, October 19–22, 2015. Communications in Computer and Information Science / Ed. by Vladimir Vishnevsky, Dmitry Kozyrev. — Springer, 2016. — P. 177–187.
2. *Anulova S.* Properties of Fluid Limit for Closed Queueing Network with Two Multi-servers // 19th International Conference "Distributed Computer and Communication Networks", Moscow, Russia, November 21–25, 2016. Communications in Computer and Information Science / Ed. by Vishnevskiy, V. M.; Samouylov, K. E. & Kozyrev, D. V. — Springer, 2016. — P. 369–380.
3. *Whitt W.* Fluid models for multi-server queues with abandonments // Oper. Res. — 2006. — Vol. 54, no. 1. — P. 37–54, http://pubsonline.informs.org/doi/abs/10.1287/opre.1050.0227