UDC 519.87

On the Rate of Convergence to stationarity of the Unreliable Queueing Network with Dynamic Routing

E. Yu. Kalimulina*

* *V. A. Trapeznikov Institute of Control Sciences,*
Russian Academy of Sciences,
Profsoyuznaya st. 65, Moscow, 11799, Russia

Abstract. In this paper we consider a Jackson type queueing network with unreliable nodes. The network consists of $m < \infty$ nodes, each node is a queueing system of M/G/1 type. The input flow is assumed to be the Poisson process with parameter $\Lambda(t)$. The routing matrix $\{r_{ij}\}$ is given, $i, j = 0, 1, ..., m$, $\sum_{i=1}^{m} r_{0i} \leq 1$. The new request is sent to the node i with the probability r_{0i}, where it is processed with the intensity rate $\mu_i(t, n_i(t))$. The intensity of service depends on both time t and the number of requests at the node $n_i(t)$. Nodes in a network may break down and repair with some intensity rates, depending on the number of already broken nodes. Failures and repairs may occur isolated or in groups simultaneously. In this paper we assumed if the node j is unavailable, the request from node i is send to the first available node with minimal distance to j, i.e. the dynamic routing protocol is considered in the case of failure of some nodes. We formulate some results on the bounds of convergence rate for such case.

Keywords: dynamic routing, queueing system M/G/1, unreliable network, Jackson network, convergence rate.

1. Introduction

Queueing systems and networks are the most suitable mathematical tools for modelling and performance evaluation of complex systems such as modern computer systems, telecommunication networks, transport, energy and others [1–3]. A large number of research papers study queueing systems with unreliable servers. [4]. The less ones consider queueing networks.

This work is motivated by a practical task of modelling of modern telecommunication networks. We propose a modification of the open queueing network model, based on the principle of dynamic routing.

There are some math research papers where queueing networks with dynamic routing were considered. Queueing networks with constant routing matrix were considered in papers [5, 6], each node there was modelled as a multichannel system, principle of dynamic routing was a random selection of a channel at the node. There are some researches on unreliable queueing networks. The common idea for modifying the routing matrix is blocking of requests and repeated service after nodes recovery. The result related to the rate of convergence to the stationary distribution for unreliable network is given in [7, 8]. In this paper we give some results for unreliable networks similarly as it was done in [8], but we propose another approach to the modification of the route matrix $\{r_{ij}\}$ and consider in a more general model for network nodes.

2. Process Definition

It is assumed that nodes at the network are unreliable and may break down or repair. Failures can be both individual and in a group (as in models in [7,8]). We will refer to $M_0 = \{0, 1, 2, ..., m\}$ as the set of nodes, where "0" is the "external node" (entry and exit from the network) and to $D \subset M$ as the subset of failed nodes, $I \subset M \setminus D$ the subset of working nodes, nodes from I may break down with the intensity $\alpha^D_{D\cup I}(n_i(t))$. Nodes from $H \subset D$ may recover with the intensity $\beta^D_{D\setminus H}(n_i(t))$. It is assumed the routing matrix (s_{ij}) is given. Additionally the adjacency matrix for our network (s_{ij}) is considered, where $s_{ij} = 1, \quad \text{if} \quad r_{ij} \neq 0$, and $s_{ij} = 0, \quad \text{if} \quad r_{ij} = 0$.

Now we can consider all possible paths of the network graph. To find them we need to calculate the following matrix: $(s_{ij})^2$, $(s_{ij})^3$, ..., $(s_{ij})^m$, $m < \infty$, $(s_{ij})^1 = (s_{ij})$. The matrix $(s_{ij})^m$ has the following property: the element in row i and column j is the number of paths from node i in the unit j of length m (including $(m-1)$ transitional nodes).

We take the following routing scheme for network nodes from the subset D (we call it as "dynamic routing without blocking"). Only transitions to $M_0 \setminus D$ are possible for nodes from D:

$$r^D_{ij} = \begin{cases} 0, \quad \text{if} \quad j \in D, i \neq j, \\ r_{ij} + r_{ik}/s^p_{ik}, \quad \text{if} \quad j \notin D, k \in D \\ \ \exists\ i \to j \to i' \to j' \to ... \to i'' \to k : \\ \underbrace{s^1_{ij} * s^1_{ji'} * s^1_{i'j'} * ... * s^1_{i''k}}_{p+1} \neq 0, \\ \text{where} \quad p = \min\{2, 3, ..., m : s^p_{ik_{k\in D}} \neq 0\}, \\ r_{ii} + \sum_{\substack{k\in D \\ s^p_{ik}=0\ \forall\ 1<p\leq m}} r_{ik}, \quad \text{if} \quad i \in M_0 \setminus D, i = j, \end{cases}$$

where s^p_{ik} - element of a matrix $(s_{ij})^p$.

The routing matrix is changed according to the same way for the input flow:

$$\Lambda r_{0j}^{D} = \begin{cases} \Lambda r_{0j}, \quad \text{if} \quad j \in M \setminus D, \\ \Lambda(r_{0j} + r_{0k}/s_{0k}^{p} * \underbrace{(s_{0j}^{1} * s_{ji'}^{1} * s_{i'j'}^{1} * ... * s_{i''k}^{1})}_{p+1}), \\ \\ \text{ if } j \notin D, k \in D \\ 0, \quad \text{otherwise.} \end{cases}$$

Further we will refer to the modified routing matrix as $R^D = (r_{ij}^D)$, the intensities of failures and recoveries depend on the state of nodes and does not depend on network load and are defined as $\alpha(D, I)$, and $\beta(D, H)$.

A more general model than in [7] is considered for network nodes. It is assumed that each network node is a queueing system type $M/G/1$. The system‘s dynamic will be described by a continuous in time random process $X(t)$ taking values from the following enlarged state space $\mathbb{E}$:

$$\tilde{\mathbf{n}} = ((n_1, z_1), (n_3, z_2), ..., (n_m, z_m), D) \in \{\mathbb{Z}_+ \times \{R_+ \cup 0\}\}^m \times |D| = \mathbb{E},$$

where n_i is the number of requests at the node i, z_i - time elapsed from the beginning of service for the current request i, $|D|$ - the cardinality of set D. Intensity rates $\mu_i(n_i, z_i)$ depend on both the number of requests at nodes $n_i(t)$ and time $z_i(t)$, time elapsed from the beginning of service for the current request at time t.

The following transitions in a network are possible:

$$\begin{aligned} T_{ij}\tilde{\mathbf{n}} &:= (D, n_1, \cdots, n_i - 1, \cdots, n_j + 1 \cdots, n_m), \\ T_{0j}\tilde{\mathbf{n}} &:= (D, n_1, \cdots, n_j + 1, \cdots, n_m), \\ T_{i0}\tilde{\mathbf{n}} &:= (D, n_1, \cdots, n_i - 1, \cdots, n_m), \\ T_{H}\tilde{\mathbf{n}} &:= (D \setminus H, n_1, \cdots, n_m), \\ T^{I}\tilde{\mathbf{n}} &:= (D \cup I, n_1, \cdots, n_m). \end{aligned}$$

Definition 1 *The Markov process $\mathbf{X} = (X(t), t \geq 0)$ is called unreliable queueing network if it's defined by the following infinitesimal generator:*

$$\tilde{\mathbf{Q}}f(\tilde{\mathbf{n}}) = \sum_{j=1}^{m}[f(T_{0j}\tilde{\mathbf{n}}) - f(\tilde{\mathbf{n}})]\Lambda(t)r_{0j}^{D}$$

$$+\sum_{i=1}^{m}\sum_{j=1}^{m}[f(T_{ij}\tilde{\mathbf{n}}) - f(\tilde{\mathbf{n}})]\mu_i(n_i(t), z_i(t))r_{ij}^{D}$$

$$+\sum_{I\subset M}[f(T^I\tilde{\mathbf{n}})-f(\tilde{\mathbf{n}})]\alpha(D,I)+\sum_{H\subset M}[f(T^I\tilde{\mathbf{n}})-f(\tilde{\mathbf{n}})]\beta(D,H)$$

$$+\sum_{j=1}^{m}[f(T_{j0}\tilde{\mathbf{n}})-f(\tilde{\mathbf{n}})]\mu_j(n_i(t),z_i(t))r_{j0}^D.$$

3. Main results

Like the classical Jackson network the existence of a stationary distribution for an unreliable network with dynamic routing may be proved.

Theorem 1 *It is assumed the following conditions for unreliable network from the Definition 1*

$$1)\ \inf_{n_j,t}\mu_j(n_j,z_j)>0 \qquad \forall\ j,$$

2) time of service and time between new arrivals are independent random variables,
3) routing matrix R^D is reversible,
then the stationary distribution for unreliable networks is defined by formulae

$$\pi(\tilde{\mathbf{n}})=\pi(D,n_1,n_2,\cdots,n_m)=\frac{1}{C}\frac{\psi(D)}{\phi(D)}\prod_{i=1}^{m}\frac{1}{C_i}\frac{\lambda_i^{n_i}}{\prod_{k=1}^{n_i}\mu_i(k)}$$

where

$$C_i=\sum_{n=0}^{\infty}\frac{\lambda_i^{n_i}}{\prod_{y=1}^{n}\mu_i(y)},\ \ \lambda_i=\sum_{j=0}^{m}\Lambda*r_{ji}.$$

The main result for the convergence rate is formulated in terms of the spectral gap for unreliable queueing network. The preliminary notations and results on the spectral gap: there is a Markov process $\mathbf{X}=(X_t,t\geq 0)$ with the matrix of transition intensities $Q=[q(\mathbf{e},\mathbf{e}')]_{e,e'\in\mathbb{E}}$, with stationary distribution π and an infinitesimal generator given by

$$\mathbf{Q}f(\mathbf{e})=\sum_{\mathbf{e}'\in\mathbb{E}}(f(\mathbf{e}')-f(\mathbf{e}))q(\mathbf{e},\mathbf{e}').$$

The usual scalar product on $L_2(\mathbb{E},\pi)$ is defined as

$$\langle f,g\rangle_{pi}=\sum_{\mathbf{e}\in\mathbb{E}}f(\mathbf{e})g(\mathbf{e})\pi(\mathbf{e}).$$

The spectral gap for $\mathbf{X}$ is

$$Gap(\mathbf{Q}) = \inf\{-\langle f, \mathbf{Q}f\rangle_\pi : \|f\|_2 = 1, \langle f, \mathbf{1}\rangle_\pi = 0\}.$$

The main result for a network is formulated in the following theorems:

Theorem 2 *If* $\mathbf{X}$ *is a Markov process with infinitesimal generator* $\mathbf{Q}$*, it is assumed that* $\mathbf{Q}$ *is bounded, the minimal intensity of service is strictly positive* $\inf_{n_j,t} \mu_j(n_j, z_j) > 0$ *and the routing matrix* (r_{ij}^D) *is reversible, then* $Gap(\mathbf{Q})$ *is strongly positive, if the following condition is true: for any* $i = 1, \cdots, m$*, for the birth and death process, corresponding to the node* i *with parameters* λ_i *and* $\mu_i(n_i, z_i)$ *the spectral gap is strictly positive* $Gap_i(\mathbf{Q}_i) > 0$*.*

Theorem 3 *If* $\mathbf{X}$ *is a Markov process with a bounded infinitesimal generator* $\mathbf{Q}$*, positive minimal intensity of service* $\inf_{n_j,t} \mu_j(n_j, z_j) > 0$ *and reversible routing matrix* (r_{ij}^D)*, then* $Gap(\mathbf{Q}) > 0$ *iff for any* $i = 1, \cdots, m$*, the distribution* $\pi = (\pi_i), i \geq 0$ *has light tails, i.e. the following condition* $\inf_k \frac{\pi_i(k)}{\sum_{j>k} \pi_i(j)} > 0$*.*

References

1. *Lakatos L., Szeidl L., Telek M.* Introduction to Queueing Systems with Telecommunication Applications. — Springer Science & Business Media, Mathematics, 2012. — 388 p.
2. *Daigle J.* Queueing Theory with Applications to Packet Telecommunication. — Springer Science & Business Media. Technology & Engineering. Jan 16, 2006. — 316 p.
3. *Thomasian A.* Analysis of Fork/Join and Related Queueing Systems // ACM Comput. Surv. — 2014. — Vol. 47, no. 2. — Article No. 17. — dx.doi.org/10.1145/2628913.
4. *Jain M., Sharma G.C., Sharma R.* Unreliable server M/G/1 queue with multi-optional services and multi-optional vacations // International Journal of Mathematics in Operational Research. — 2013. — Vol. 5, no. 2. — P. 145–169.
5. *Vvedenskaya N.D.* Configuration of overloaded servers with dynamic routing // Probl. Inf. Transm. — 2011. — Vol. 47, no. 3. — P. 289–303.
6. *Sukhov Yu.M., Vvedenskaya N.D.* Fast Jackson Networks with Dynamic Routing // Prob. Inf. Transm. — 2002. — Vol. 38, no. 2. — P. 136–153.
7. *Lorek P., Szekli R.*, Computable bounds on the spectral gap for unreliable Jackson networks. // Adv. in Appl. Probab — 2015. — Vol. 47. — P. 402–424.
8. *Lorek P.* The exact asymptotic for the stationary distribution of some unreliable systems // arXiv:1102.4707 [math.PR]. 2011.
9. *Chen M.F.* Eigenvalues, Inequalities, and Ergodic Theory. — Springer, 2005.