EFFICIENCY ANALYSIS OF STOCHASTIC MODEL VALIDATION BY USE OF TRACE-DRIVEN SIMULATION

Alexander Andronov
Department of Mathematical Support of
Transport System Control
Riga Technical University
Kalku street 1, LV-1658, Riga, Latvia
E-mail: lora@mailbox.riga.lv

KEYWORDS

Model validation, Trace-driven simulation, power function, semi-Markov process

ABSTRACT

The "Trace-driven simulation" is often employed in practical problems for the validation of a considered stochastic model. In this paper we propose a quite general model for description of this method by use of embedded semi-Markov process. It allows us investigating its efficiency.

INTRODUCTION

"Trace-driven simulation" is an approach that is used for validation of considered model (Kleijnen et al. 1999). This approach can be described by the following way. We suppose that a real system of interest has observed input $A_1, A_2, ..., A_n$ and output or replicate $W_1, W_2, ..., W_m$. For example, $A_1, A_2, ..., A_n$ form a sequence of arrival times of customers into a queuing system, $W_1, W_2, ..., W_m$ form a sequence of sojourn times of the customers in the system or the output times from the system. The output series are evaluated through a performance measure (response) *X*, the average sojourn time of the customers, for example.

We have a simulated system as well. We should verify its adequacy to the real system: it is the so-called null hypothesis. For this purpose we use the same "trace" $A_1, A_2, ..., A_n$ and produce the simulation *B* times. As a result we get *B* simulated outputs (replicates) $\widetilde{W}_1^{(r)}, \widetilde{W}_2^{(r)}, ..., \widetilde{W}_m^{(r)}$, where r is a run number of the simulation, $r = 1,2,...,B$. Let $\widetilde{X}^{(1)}, \widetilde{X}^{(2)}, ..., \widetilde{X}^{(B)}$ be a corresponding sequence of the performances (responses).

To validate the simulation model (to test the null-hypothesis), the real and simulated performances X and $\widetilde{X}^{(1)}, \widetilde{X}^{(2)}, ..., \widetilde{X}^{(B)}$ should be compared statistically. Some solutions of this problem were given in (Kleijnen et al. 1999).

We consider another approach. The simplest way supposes a sorting of the simulated performances $\widetilde{X}^{(1)}, \widetilde{X}^{(2)}, ..., \widetilde{X}^{(B)}$. It gives the order statistics $\widetilde{X}_{(1)}, \widetilde{X}_{(2)}, ..., \widetilde{X}_{(B)}$ and the estimated α-quantile of distribution $\widetilde{X}_{(\lfloor B\alpha \rfloor)}$. This procedure gives, for example, a two-sided ($1-\alpha$)-confidence interval for the original value X, ranging from the lower estimated $\alpha/2$-quintile to the upper $(1-\alpha/2)$-quintile. If X value falls outside this interval then we reject the simulated model (the null-hypothesis). In addition, the significance level of this testing is equal to α.

Our aim is to investigate the efficiency of the described approach for quite general case of complex system. We shall consider these systems as composite two-components embedded semi-Markov process.

The next section presents the mathematical model of considered systems. Section 3 describes a validation procedure. Section 4 is devoted to a calculation of a power function for proposed approach. A special case of queueing system $GI/M/S$ is presented in section 5.The last section contains some concluding remarks.

MATHEMATICAL MODEL

We shall suppose that observed input $\{A_\nu\}$ corresponds to a semi-Markov process $A(t)$. The last is described by discrete set ε_A of states, matrix of one-step transition probabilities $\boldsymbol{P} = (P_{i,j})$ and the distribution function $F_i(t)$ of sojourn time in each state $i \in \varepsilon_A$. Obviously,

$$\sum_{j \in \varepsilon_A} P_{i,j} = 1, \quad F_i(\infty) = 1 \quad \forall \ i \in \varepsilon_A .$$

Remind (Ross 1992, p.86) that "…semi-Markov process $(A(t))$ records the state of the process at each time point …". So we have these states and their durations as considered input.

Proceedings 12th International Conference ASMTA 2005
Khalid Al-Begain, Gunter Bolch, Miklos Telek © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

Further we suppose that the inner structure of our simulated system is described by a stochastic process $S(t)$ with discrete set of states ε_S. If state $J \in \varepsilon_A$ of Semi-Markov process $A(t)$ takes place then process $S(t)$ behaves as independent Semi-Markov process with matrix $\boldsymbol{P}^{S,J}$ of one-step transition probabilities and distribution function $F_i^{S,J}(\cdot)$ of sojourn time in the state $i \in \varepsilon_S$.

We shall name jumps (transition) of process $A(t)$ as jumps of the first type and supplemental jumps of the process $S(t)$ only as the jumps of the second type. Note that the probability having a jump of the process $A(t)$ at the time t does not depend on state of $S(t)$.

Let t^* be a time moment when a jump of the first type takes place, $J_- = A(t^*-)$ and $J_+ = A(t^*+)$ be states of the process $A(t)$ immediately before and after this time moment. Let $J_-{}^S = S(t^*-)$ and $J_+{}^S = S(t^*+)$ be the same for the process $S(t)$. If J_+ is entered, the next state J_+^S of the process $S(t)$ is chosen according to the transition probabilities $q_{J_-^S,\, J_+^S}^{J_-,\, J_+}$. If the state J_+^S is chosen, then the time until the next transition of process $S(t)$ has distribution function $F_{J_+^S}^{S,J_+}$ as usually for $S(t)$.

Now we consider sequence of time moments $t_1, t_2, \ldots$ when all jumps occur. Every one of such moments can produce random variable $\tilde{W}_1, \tilde{W}_2, \ldots$ − output of our system. To describe corresponding stochastic mechanism, we introduce the following notations for the jump of the first type at the time moment t^*: $\tau(t^*)$ − length of an interval between the previous and current jumps of the first type; $U(t^*)$ − the time since the last jump of the processes $S(t)$ till t^*, so-called *age* of current state at t^*. Then the probabilities $\Pr_1(J_-, J_+, J_-{}^S, J_+{}^S, \tau(t^*), U(t^*))$ to produce the random variables $\tilde{W}$ are functions of $J_-, J_+, J_-{}^S, J_+{}^S, \tau(t^*)$ and $U(t^*)$. For the jumps of the second type this probability is denoted by $Pr_2(J, J_-{}^S, J_+{}^S, U(t^*))$. If the corresponding random variable $\tilde{W}$ is produced, then its distribution function is given by formulas

$$G_1{}^S(x; J_-, J_+, J_-{}^S, J_+{}^S, \tau(t^*), U(t^*)) =$$
$$= P\{\tilde{W} \le x / J_-, J_+, J_-{}^S, J_+{}^S, \tau(t^*), U(t^*)\},$$
$$G_2{}^S(x; J, J_-{}^S, J_+{}^S, U(t^*)) =$$
$$= P\{\tilde{W} \le x / J, J_-{}^S, J_+{}^S, U(t^*)\}.$$

VALIDATION PROCEDURE

The sequence $\{\tilde{W}_i\}$ is an output from the simulated system with the inner structure that is described by the process $S(t)$. Let the inner structure of the real system is described by random process $B(t)$. The null-hypothesis supposes that process $B(t)$ is similar to process $S(t)$. Let $W_1, W_2, \ldots$ denote the output for real process $B(t)$ if the same input $\boldsymbol{A} = (A_1, A_2, \ldots, A_n)$ takes place. We must compare two sequences $\{W_i\}$ and $\{\tilde{W}_i\}$ statistically. For this purpose we use some performance measure $X = \xi(W_1, W_2, \ldots)$ and $\tilde{X} = \xi(\tilde{W}_1, \tilde{W}_2, \ldots)$.

Let us calculate a conditional distribution function $R_S(x \mid \boldsymbol{A})$ of $\tilde{X}$ by condition that input (trace) $\boldsymbol{A}$ is fixed:

$$R_S(x \mid \boldsymbol{A}) = P\{\tilde{X} \le x \mid \boldsymbol{A}\}.$$

Let $R_S{}^{-1}(\gamma \mid \boldsymbol{A})$ be the corresponding γ-quantile: $R_S(R_S{}^{-1}(\gamma \mid \boldsymbol{A}) \mid \boldsymbol{A}) = \gamma$, $R_B(x \mid \boldsymbol{A}) = P\{X \le x \mid \boldsymbol{A}\}$ - be corresponding notation for the random variable X and process $B(t)$ instead of $R_S(x \mid \boldsymbol{A})$ for $\tilde{X}$ and $S(t)$.

Now we are able to test the null hypothesis H_0 that random processes $B(t)$ and $S(t)$ have the same probabilistic structure against the alternative hypothesis H_1 that they have different structures. Let us suppose (without restriction of generality) that if the alternative hypothesis H_1 is true then the performance measure X is stochastically greater than $\tilde{X}$. In this case we reject the null-hypothesis for the significance level α if fixed value X exceeds $R_S^{-1}(1-\alpha \mid \boldsymbol{A})$. In other words a decision interval for the null-hypothesis H_0 is $(0, R_S^{-1}(1-\alpha \mid \boldsymbol{A}))$.

If hypothesis H_1 is true, the probability that X falls outside the interval $(0, R_S^{-1}(1-\alpha \mid \boldsymbol{A}))$ determines value of a power function:

$$\beta(H_1 \mid \boldsymbol{A}) = P_B\{X > R_S^{-1}(1-\alpha \mid \boldsymbol{A}) \mid \boldsymbol{A}\} =$$
$$= 1 - R_B(R_S^{-1}(1-\alpha \mid \boldsymbol{A}) \mid \boldsymbol{A}). \qquad (1)$$

Also the conditional power function $\beta(H_1 \mid \boldsymbol{A})$ for fixed input $\boldsymbol{A}$ may be calculated. Note that often corresponding calculations can be produced analytically, numerically or by simulation. We use the simulation as it was described in the Introduction.

To get the unconditional power function $\beta(H_1)$ we have to use averaging in accordance with the distribution of the input $\boldsymbol{A}$ (trace). The simplest method of this averaging consists in Monte-Carlo use.

Also, we produce k realizations of Monte-Carlo. In the l-th realization $(l = 1,2,...,k)$ we generate input $\boldsymbol{A}(l) = (A_1(l), A_2(l), ..., A_{n(l)}(l))$ and calculate the power function $\beta(H_1 \mid \boldsymbol{A}(l))$ as it was described above. Now we are able to average given results to get estimator

$$\beta^0(H_1) = \frac{1}{k} \sum_{l=1}^{k} \beta(H_1 \mid \boldsymbol{A}(l)) \qquad (2)$$

of unconditional power function $\beta(H_1)$. The law of Large Numbers states that with probability one $\beta^0(H_1)$ tends to $\beta(H_1)$ when k tends to the infinity.

CALCULATION OF THE CONDITIONAL DISTRIBUTION FUNCTION OF $\widetilde{X}$

A subject of our consideration is numerical calculation of the conditional distribution function $R_S(x \mid \boldsymbol{A})$ of the performance measure $\widetilde{X}$. The distribution function $R_B(x \mid \boldsymbol{A})$ of X has a similar presentation. Therefore it allows us calculating the power function (1). Below we shall suppose that performance measures $\widetilde{X}$ and X has additional structure, namely $\widetilde{X} = \widetilde{W}_1 + \widetilde{W}_2 + ...$ and $X = W_1 + W_2 + ...$.

Then we consider the random process $(A(t), S(t))$. Let $((J_1, T_1), (J_2, T_2), ... (J_n, T_n))$ be a sequence of states and sojourn times in these states that has been registered in the given realization of $A(t)$. We denoted it by $\boldsymbol{A}$ and named "trace".

Note that our process $(A(t), S(t))$ produces an embedded Markov chain for the time moments $C_0 = 0,\ C_1 = T_1,\ \ C_2 = C_1 + T_2, \ldots,\ \ C_n = C_{n-1} + T_n$.

Our aim is to calculate a conditional distribution of random variables (Y_ν, Z_ν), $\nu = 1,2,...$, by condition that trace $\boldsymbol{A}$ is fixed and $Y_0 = k_0,\ Z_0 = 0$. Here Y_ν denotes a state $S(C_\nu +)$ immediately after time moment C_ν, Z_ν denotes a cumulative sum of $\widetilde{W}$ for this moment. We denote corresponding conditional distribution function

$$R_\nu(k, z \mid \boldsymbol{A}) = P\{Y_\nu = k,\ Z_\nu \le z \mid \boldsymbol{A}\}.$$

Let $Q_{i,j}(u, z; t, J)$ be a probability that semi-Markov process $S(\cdot)$ with characteristics $\boldsymbol{P}^{S,J}$ and $\{F_i^{S,J},\ i \in \varepsilon_S\}$ goes from state $i \in \varepsilon_S$ to state $j \in \varepsilon_S$ at the time t, its cumulative value of Z will be less than z, age $U(t)$ of state j at the time t will be less then u:

$$Q_{i,j}(u, z; t, J) = P\{S(t) = j,$$
$$U(t) \le u,\ Z(t) \le z \mid A(\zeta)_1 =$$
$$= J,\ 0 \le \zeta \le t;\ S(0) = i, U(0) = 0, Z(0) = 0\}.$$

Note that values of $Q_{i,j}$ are calculated on the base of $Pr_2(\cdot)$ and $G_2^S(\cdot)$.
The technique of these transition probability calculation is well known (Ross 1992) and we regard ones are known.
Usual Markov type relationships give us the following expressions:

$$R_0(k, z \mid \boldsymbol{A}) = \begin{cases} 1 & if \quad k = k_0,\ z > 0, \\ 0 & otherwise. \end{cases} \qquad (3)$$

$$R_{\nu+1}(k, z \mid \boldsymbol{A}) = \sum_{i \in \varepsilon_S} \int_{-\infty}^{\infty} R_\nu(i, d\zeta \mid \boldsymbol{A}) \sum_{j \in \varepsilon_S} q_{j,k}^{J_\nu, J_{\nu+1}} \cdot$$
$$\left\{ \int_0^{T_\nu} \left[1 - Pr_1\left(J_\nu, J_{\nu+1}, j, k, T_\nu, u\right)\right] \cdot \right. \qquad (4)$$
$$\cdot Q_{i,j}\left(du, z - \zeta; T_\nu, J_\nu\right) + \int_0^{T_\nu} \Pr_1\left(J_\nu, J_{\nu+1}, j, k, T_\nu, u\right) \cdot$$
$$\cdot \int_{-\infty}^{\infty} G_1^S(dw; J_\nu, J_{\nu+1}, j, k, T_\nu, u),$$
$$\left. Q_{i,j}\left(du, z - \zeta - w; T_\nu, J_\nu \right)\right\},$$
$$k \in \varepsilon_S,\ \ \forall z \in (-\infty, \infty),\ \nu = 0,1,...,n-1.$$

Finally we have

$$R_S(x \mid \boldsymbol{A}) = \sum_{k \in \varepsilon_S} R_n(k, x \mid \boldsymbol{A}). \qquad (5)$$

It should be noted that in the last formula the expression for R_n could be modified to take into account the end of the considered process.

SPECIAL CASE

We consider queueing system $GI/M/S$: an arbitrary renewal process of customer arrivals, exponential distributed service time and the number of servers equals S. We suppose that originally system is empty. We observe a number n of customer arrivals, the sequence $C_1 = 0, C_2, ..., C_n$ of arrival times and the sequence $V_1, V_2, ..., V_n$ of their output times from the system. As performance measure we have common sojourn time of all arrived customers

$$\tilde{X} = \sum_{i=1}^{n} (V_i - C_i).$$

Our aim is to verify statistical hypothesis (null-hypothesis H_0) that service time has exponential distribution with parameter λ_0 against alternative hypothesis H_1 that this parameter equals $\lambda < \lambda_0$.

In this case it is convenient to consider only one state of process $A(t)$: $\varepsilon_A = \{1\}$. Then the observed input $\{A_\nu\}$ is a Renewal Process $T_1, T_2, ..., T_n$ (Ross 1992), where $T_1 = C_1$, $T_\nu = C_\nu - C_{\nu-1},\ \nu = 2,3,...,n.$ Let the stochastic process $S(t)$ denote the number of customers in the system at time moment t. $S(t)$ is a death process for $C_\nu < t < C_{\nu+1}$ so $P^{S,1}_{i,i-1} = 1$ for matrix $\boldsymbol{P}^{S,1}$, moreover for $t > 0$

$$F_i^{S,1}(t) = \begin{cases} 1 - e^{-\lambda_0 it}, & 0 \le i \le S, \\ 1 - e^{-\lambda_0 St}, & i > S. \end{cases}.$$

Here the jumps of the first and the second type correspond to customer arrivals and outputs from the system. Therefore

$$q^{1,1}_{i,i+1} = 1, \ \ Pr_1(\cdot) = 0, \ \ Pr_2(\cdot) = 1.$$

To use our general results for $\tilde{X} = \tilde{W}_1 + \tilde{W}_2 + ... + \tilde{W}_n$ calculation we are able to consider new "output" $\tilde{W}_\nu = V_\nu - C_\nu$. It was shown in Appendix that for $j \le i \le S,\ t \le T_\nu$

$$Q_{i,j}(T_\nu, z; t, 1) = \binom{i}{j}\left(1 - e^{-\lambda_0 t}\right)^{i-j} \cdot$$
$$\cdot e^{-\lambda_0 jt} F_{i-j}(z - jt), \ jt \le z \le it,$$

where

$$F_m(x) = (1 - e^{-\lambda_0 t})^{-m} \sum_{\nu=0}^{\nu(x)} (-1)^\nu \binom{m}{\nu} \cdot$$
$$\cdot \left[e^{-\lambda_0 \nu t} - e^{-\lambda_0 x} \sum_{\zeta=0}^{m-1} \lambda_0^\zeta \frac{(x - \nu t)^\zeta}{\zeta!} \right],$$

$0 \le x \le mt, \nu(x) = \min\{m-1, \lfloor x/t \rfloor\}$, $\lfloor x/t \rfloor$ - whole part of x/t.

The analogous formulas for the cases $i \ge j \ge S$ and $j < S \le i$ are given in the Appendix.

Now we are able to perform the calculation in accordance with the formulas (1) – (4). If after the last arrival the system contains k customers, then their rest sojourn time has Erlang distribution with parameters k and λ_0. Therefore the distribution function (5) of the total sojourn time

$$R_S(x \mid \boldsymbol{A}) = \sum_{k=1}^{n} \int_0^x R_n(k, x - z \mid \boldsymbol{A}) \cdot \cdot \lambda_0 (\lambda_0 z)^{k-1} \cdot \frac{1}{(k-1)!} \cdot e^{-\lambda_0 z}, \ x \ge 0. \qquad (6)$$

Note that the queueing systems $GI/G/\infty$ and $GI/M/S$ were considered before by use of quite complex algebraic computation (Andronov 2000, 2001). Our numerical approach allows avoiding such difficulties.

NUMERICAL EXAMPLE

Let us consider the previous case for $S = 1$. Therefore we have a single queueing system. It allows us to simplify the general formulas:

$$R_1(k, z \mid \boldsymbol{A}) = \begin{cases} 1 & if \ \ k = 1, \ z > 0, \\ 0 & otherwise, \end{cases}$$

$$R_{\nu+1}(k, z \mid \boldsymbol{A}) = \sum_{i=k-1}^{\nu} \int_0^z R_\nu(i, z - x)\, Q_{i,k-1}(T_\nu, dx; T_\nu, 1).$$

The formulas for $Q_{i,j}$ calculation (see Appendix) have the following expressions. Since $S = 1$ the case 2) $j \le i \le S$ is absent (it is covered by other cases). In the case 1) $i \ge j \ge 1$ the formula (6) takes place for $S = 1$.

In the case 3) $j = 0 < 1 \le i$, instead of (8) we have the formula:

$$Q_{i,0}(t,z;t,1) = \lambda_0 \int_0^t Q_{i,1}(\tau,z;\tau,1)d\tau .$$

Let us consider the following numerical data. The number n of considered customer is equal to 3, the first customer arrives into the empty system ($C_1 = 0$), the next customers arrive at the time moments $C_2 = 0.5$ and $C_3 = 1.5$. Our null hypothesis H_0 is $\lambda = \lambda_0 = 1$. Let the significance level $\alpha = 0.1$.

As a criterion of the hypothesis testing we use the total sojourn time of all customers. Their distribution function is defined by formula (6). The Table 1 contains its numerical values. We see that $R_S^{-1}(1-\alpha \mid \boldsymbol{A}) = R_S^{-1}(0.9 \mid \boldsymbol{A}) \approx 6$. The Table 2 contains values of the power function $\beta(\lambda)$ for the alternative hypothesis $\lambda < 1$. We see that the Trace-driven simulation allows quickly detecting a deviation from the null hypothesis.

Table 1. The conditional distribution function $R_S(x \mid \boldsymbol{A})$ under the hypothesis H_0: $\lambda = 1$

x	3	4	5	6	7
$R_S(x \mid \boldsymbol{A})$	0.487	0.671	0.809	0.897	0.948

x	8	9	10	11
$R_S(x \mid \boldsymbol{A})$	0.974	0.988	0.994	0.998

Table 2. The power function $\beta(\lambda)$ for the alternatives H_1: $\lambda < 1$

λ	1	0.9	0.8	0.7	0.6
$\beta(\lambda)$	0.103	0.146	0.205	0.283	0.382

λ	0.5	0.4	0.3	0.2	0.1
$\beta(\lambda)$	0.503	0.641	0.783	0.907	0.983

CONCLUDING REMARKS

We considered the quite general mathematical model for description of Trace-driven simulation that is used for stochastic model validation. This approach is based on the theory of embedded two-dimensional semi-Markov processes. This allows us to use proposed numerical algorithm for the efficiency investigation of the Trace-driven simulation. The future research will deal with applications of considered approach to different practical problems of stochastic model validation.

REFERENCES

Andronov, A. 2000. "Validation testing of traffic models using Trace-driven Simulation". In *Proceedings of Nordic-Baltic Transport Research Conference.* Volume 1. (Riga, Latvia, April 13-14). NTF, Riga, 1–5.

Andronov, A. 2000. "On efficiency of Trace-driven simulation". In *Proceedings of Second International Conference Simulation, Gaming, Training and Business Process Reengineering in Operations.* (Riga, Latvia, September 8-9). Riga Technical University, Riga, 107–111.

Andronov, A. 2001. "Validation testing of queueing model G/M/S by use of Trace-driven simulation". In *Proceedings of the Sixteenth Belarussian Winter Workshop on Queueing Theory.* (Minsk, Belarus, January 23-25). Belarus State University, Minsk.

Feller, W. 1971. *An Introduction to Probability Theory and its Applications*. Volume II. John Wiley &Sons, New York.

Gentle, J.E. 2002. *Elements of Computational Statistic*. Springer-Verlag, New York-Berlin-Heildelberg.

Kleijnen, J.P.C.; B.Bettonvil; and W.Van Groendaal. 1998. "Validation of Trace-driven simulation models: a novel regression test". *Management Science,* 44, 812–819.

Kleijnen, J.P.C.; R.C.H.Cheng; and B.Bettonvil. 1999. *Validation of Trace-driven models: bootstrap tests.* Working Paper, Department of Information Systems, Tilburg University, The Netherlands.

Ross, Sheldon M. 1992. *Applied probability models with optimization applications.* Dover Publications, INC., New York.

APPENDIX

Let us derive formulas for the calculation of $P_{i,j}(t,x) = Q_{i,j}(T_V, x; t, 1)$. It means that we have queuing system G/M/S with S servers. Service time has exponential distribution with parameter λ. We suppose that i customers are in the system at the beginning and new customers do not arrive during time t. Then $P_{i,j}(t,x)$ is probability that j customers will be in the system at time t and total sojourn time of all customers in the system during time t will be less than x. The probability $P_{i,j}(t,x)$ depends on values i and j. Let us consider three possible cases separately.

1) $i \ge j \ge S$. In this case S customers are served constantly. The total service intensity is equal to λS. The output flow of served customers is the Poisson process with this intensity. Therefore, probability that $i - j$ customers will be served during the time t equals

$$\frac{1}{(i-j)!}(\lambda St)^{i-j}e^{-\lambda St}, \quad S \le j \le i.$$

Now we use the following useful statement (Ross 1992, p.18): "under the condition that m events of the Poisson process have occurred in (0, t), the times at which events occur, considered as unordered random variables, are distributed independently and uniformly in the interval (0, t)." Therefore the total sojourn time of all customers in the system during time t is equal to sum of $m = i - j$ random variables that are distributed independently and uniformly in the interval (0, t) plus jt (sojourn time of j remaining customers). The distribution function of the sum is (Feller 1971):

$$U_m(x) = \begin{cases} 0, & \text{for } x < 0, \quad m = 0, \\ 1, & \text{for } x \ge 0, \quad m = 0, \\ \frac{1}{t^m m!}\sum_{\nu=0}^{m}(-1)^\nu \binom{m}{\nu}(x-\nu t)_+^m, & x \ge 0, \ m > 0, \end{cases}$$

where

$$(x-\nu t)_+ = \begin{cases} 0 & \text{if } x < \nu t, \\ x - \nu t & \text{if } x \ge \nu t. \end{cases}$$

Therefore

$$P_{i,j}(t,x) = \frac{1}{(i-j)!}(\lambda St)^{i-j}e^{-\lambda St}U_{i-j}(x-jt), \quad (7)$$
$$S \le j \le i, \quad x \ge jt.$$

2) $j \le i \le S$. In this case a number of served customers during the time t is a number of success in the Bernoulli trials for i trials and the success probability $1-\exp(-\lambda t)$. Therefore, probability to complete a service of $i - j$ customers during the time t is equal to

$$\binom{i}{j}[(1-\exp(-\lambda t))]^{i-j}\exp(-j\lambda t), 0 \le j \le i \le S.$$

Therefore the total sojourn time of all customers in the system during time t is equal to jt plus sum of $m = i - j$ independent random variables that have truncated in t on the right exponential distribution with parameter λ. It can show that the distribution functions and probability density function of this sum are given by formulas, correspondingly:

$$F_m(x) = \begin{cases} 0, & x < 0, \ m = 0, \\ 1, & x \ge 0, \ m = 0, \\ (1-e^{-\lambda t})^{-m}\sum_{\nu=0}^{\nu(x)}(-1)^\nu\binom{m}{\nu}\cdot & 0 \le x \le mt, m > 0, \\ \cdot\left[e^{-\lambda\nu t} - e^{-\lambda x}\sum_{\xi=0}^{m-1}\lambda^\xi\frac{(x-\nu t)^\xi}{\xi!}\right], & \end{cases}$$

where $\nu(x) = \min\ \{m-1, \lfloor x/t \rfloor\}$, $\lfloor x/t \rfloor$ - whole part of x/t,

$$f_m(x) = \frac{1}{(m-1)!}\lambda^m e^{-\lambda x}(1-e^{-\lambda t})^{-m}\cdot$$
$$\sum_{\nu=0}^{m-1}(-1)^\nu\binom{m}{\nu}(x-\nu t)_+^{m-1}, \ 0 \le x \le mt.$$

Finally

$$P_{i,j}(t,x) = \binom{i}{j}[1-\exp(-\lambda t)]^{i-j}$$
$$\cdot\exp(-j\lambda t)F_{i-j}(x-jt), \ j \le i \le S, \ jt \le x \le it.$$

3) $j < S \le i$. In this case there exists some time moment u, $0 < u < t$, that the number of customers in the system at u changes from S to $S-1$. The probability that S customers are in the system and the total sojourn time is less than $x - y$ at the time moment t, is equal to $P_{i,S}(u, x-y)$. The probability of service completion during interval $(u, u+du)$ is equal to $\lambda Sdu + 0(du)$. The probability, that after residual time $t-u$ we will have j customers in the system and the total sojourn time will increase less than on y, is equal to $P_{S-1,j}(t-u, y) + 0(dy)$. As the moment u and the increment y are random variables then

$$P_{i,j}(t,x) = \int_0^t \int_{(t-u)j}^{\Delta(u)} P_{i,S}(u,x-y)\cdot$$
$$P_{S-1,j}(t-u,dy)\lambda Sdu = \int_0^t \int_{(t-u)j}^{\Delta(u)} P_{i,S}(u,x-y)\binom{S-1}{j}\cdot$$
$$[1-\exp(-\lambda(t-u))]^{S-1-j}\cdot \quad (8)$$
$$\cdot\exp(-j\lambda(t-u))f_{S-1-j}(y-(t-u)j)\,\lambda S \ dy \ du,$$
$$j < S \le i, \quad x \ge 0,$$

where $\Delta(u) = \min\ \{x - uS, (S-1)(t-u)\}$.