DISCRETE MARKOV SIMULATIONS

Salah M. Yousif
California State University, Sacramento
Sacramento, California

ABSTRACT

A large number of engineering and business systems can be represented by Markov-Kolmogrov equations representing their probability transition behavior. These systems are characterized by two stochastic processes: arrival and service processes. The steady state behavior of these systems leads to Kolmogrov balance equations that lead to lucid simulation techniques.

I. INTRODUCTION

The first step in constructing a mathematical model to represent the behavior of a stochastic system is idealization. Idealization involves assumings, approximating and throwing away necessary information to make the model mathematically tractable and a good approximation to the actual system. In many stochastic system modelling, especially when the events of the system occur discreetly at random times, the Markov exponential model can be considered as the best one to represent this system. The system events occur through arrival and leave through service. (13). The arrivals are assumed Poisson distributed with interoccurrent or interarrival times having exponential distribution. The service times are assumed to have, also, a negative exponential distribution with different rates. Examples of this model are congestion systems, reliability systems, telephone exchange systems, birth and death process of a population, traffic systems, machine service, etc.

The model is defined at any time by the probability of the state and occurrence and service rates at that time. The time during the change from state to another is random with negative exponential distribution. The probability density functions of continuous exponentially distributed random variable X, with parameter λ, is given by:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geqslant 0 \\ 0 & \text{for } x < 0 \end{cases} \qquad 1.1$$

The corresponding distribution function of this random variable is given by:

$$F(x) = \begin{cases} 1-e^{-\lambda x} & \text{for } x \geqslant 0 \\ 0 & \text{for } x < 0 \end{cases} \qquad 1.2$$

The mean and the variance of the exponential distribution are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$ respectively. An important property of exponentially distributed random variable is that of being memoryless. That is,

$$P\{x>t_1 + t_2 \mid x>t_1\} = P\{x>t_2\} \qquad 1.3$$

for all t_1 and t_2. As it was mentioned earlier, the events occur or arrive with Poisson distribution with rate λ. Therefore the arrival process $\{x(t), t\geqslant 0\}$ is Poisson process with mean λt and,

$$P\{x(t_o + t) - x(t_o) = \kappa\} = e^{-\lambda t}\frac{(\lambda t)^\kappa}{\kappa!} \qquad 1.4$$

for $\kappa = 0,1,2,$ and all t_o and t. The exponential Markov stochastic model can be described mathematically by a Markov Process $\{x(t), t\geqslant 0\}$ as follows. It will be assumed, for the sake of applications, that the process $\{x(t), t\geqslant 0\}$ is homogeneous or stationary. That is the probability $P\{x(t_o + t) = r \mid x(t_o) = j\}$ is independent of t_o; $t_o < t$. The random variable $x(t_1)$ at any time t_1 will be called the state of the process.

The state $x(t)$ of the process is discrete. This means that the change from state to state occurs in jumps. Also, it will be assumed, but not exclusively, that the state ranges over the positive integer numbers which makes the process $\{x(t), t\geqslant 0\}$ a Markov counting process. The time duration spent by the process in any state before moving into another state is, by the Markov property, random variable with negative exponential distribution. The change in the process from any state i at time s to state j at time t, t>s, is described by the transition probability

$$P_{ij}(s,t) = P\{x(t) = j \mid x(s) = i\} \qquad 1.5$$

which will equate $P_{ij}(s-t)$ due to the stationarity of the process assumed earlier. Thus, the above equation may be written as

$$P_{ij}(s + t) = P\{x(s + t) = j \mid x(0) = i\} \qquad 1.6$$

The Markov transition probability $P_{ij}(t)$ satisfies all t, the following obvious conditions

$$P_{ij}(t) \geqslant 0 \quad , \lim_{t\to 0} P_{ij}(t) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad 1.6$$

$$\sum_j P_{ij}(t) = 1 \qquad 1.7$$

where j in the summation ranges over all the states of the process which could be finite or infinite. The condition , $\lim_{t\to 0} P_{ij}(t) = 1$, implies instantaneous change or jump is not permissible. This means that there is no possibility of leaving the state of the process at the moment of arrival there. The number of events occurring in the interval $[s, s + t]$ has Poisson's distribution with mean λt. That is, for any s and t

$$P_{ij}(t) = P\{x(s+t) = j \mid x(s) = i\} = e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} \qquad 1.8$$

The transition probability $P_{0n}(t) = P\{x(t)=n \mid x(0) = 0\}$ is the unconditional probability of being at state n at time t and will be denoted by $P_n(t)$. $P_n(t)$ satisfies the following condition

$$P_n(0) = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n > 0 \end{cases} \qquad 1.9$$

The static behavior of the transition for any state i at time 0 to state j at time t + s, t 0, is described by the time-continuous Chapman-Kolmogrov equations

$$P_{ij}(t + s) = \sum_k P_{ik}(s) \; P_{kj}(t), \text{ all i and j.} \qquad 1.10$$

where k is any intermediate state in the path of transition from state i to state j.

The process $x(t)$, in our model, will represent the number of occurrences of an event in time interval (0,t). It may, for example, be the number of people in the line in queueing system, the number of machine failures, the number of immigrants into a territory, and many other examples of similar nature. If we let Z_{n+1} and Y_n denote the number of arrivals and services, respectively, in the interval (t_n, t_{n+1}) then the sequence of random variables $\{X_n \; n = 0,1,2, \text{ or } 0\}$ satisfies the following relation

$$X_{n+1} = (X_n - Y_n)^+ + Z_{n+1} \qquad 1.11$$

where $(\alpha)^+ = \max(\alpha,0)$ and X_n denotes the state of the process at time t_n. If $\{X_{n+1}\}$ and $\{Y_n\}$ are sequences of independent random variables, then, the sequence $\{X_n, n = 0,1,2,\ldots\}$ will represent a discrete Markov process. In a queueing mode X_n defines the number of customers in the system at time $t_n + 0$. If the system has finite capacity k, the above relation becomes

$$X_{n+1} = \min\left\{k, (X_n + Z_n - Y_{n+1})^{+}\right\}, \quad n = 0,1,2 \qquad 1.12$$

The dynamic behavior of infinitismal transition from time t from state i into stage j leads to the Kolmogorov differential equations in the transition probability $P_{ij}(t)$.

II. DYNAMIC STATE TRANSITION

The infinitismal transition scheme which represents the dynamic discrete-state Markov process can be best understood in the light of the change in transition probability in very small time Δt. We will assume that $P_{ij}(t)$ to have derivative $P'_{ij}(t)$ for all $t>0$.

Also, we will assume the following two propositions.

$$q_i = \lim_{\Delta t \to 0} \frac{1 - P_{ii}(\Delta t)}{\Delta t} = -P'_{ii}(0), \quad q_i > 0 \qquad 2.1$$

$$q_{ij} = \lim_{\Delta t \to 0} \frac{P_{ij}(\Delta t)}{\Delta t} = P'_{ij}(0), \quad i \neq j \qquad 2.2$$

The above two propositions are actually characteristics of the exponential Markov process and their proof is quite simple. Also, as it is seen,these two propositions have the following physical and intuitive interpretation. The first proposition means that there will be transition out of state i in infinitismal time. The second says that if this transition is into state j, it will occur with rate q_{ij}. The relation between q_i and q_{ij} is given by

$$\sum_{j \neq i} q_{ij} = q_i, \quad q_{ii} = -q_i \qquad 2.3$$

The above relation can be easily proved using (1.1).

Using (2.1) and (2.2) and differentiating Chapman-Kolmogorov (5) equation with respect to each variable will yield the following two difference differential equations,

$$P'_{ij}(t) = -q_j P_{ij}(t) + \sum_{k \neq j} q_{kj} P_{ik}(t), \quad i,j = 0,1,2, N \qquad 2.4$$

$$P'_{ij}(t) = -q_i P_{ij}(t) + \sum_{k \neq i} q_{ik} P_{kj}(t), \quad i,j = 0,1,2 \ldots, N \qquad 2.5$$

The first equation for a fixed i and the second equation for fixed j, are called forward and backward Kolmogorov equations, respectively. The initial conditions for both equations are given by,

$$P_{ij}(0) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad 2.6$$

The forward and backward Kolmogorov equations can be expressed, using (2.3), in the following matrix forms, respectively,

$$P'(t) = P(t)\, Q \qquad 2.4'$$

and $$P'(t) = Q\, P(t) \qquad 2.5'$$

The initial conditions for both equations are

$$P(0) = I \qquad 2.6'$$

where $P(t)$ is the $(N+1) \times (N+1)$ probability transition matrix whose ij entry is the ij-th transition probability $P_{ij}(t)$, Q is the transition rate matrix whose ij entry is the ij-th rate q_{ij} given by (2.3); i.e.,

$$Q = \begin{bmatrix} -q_o & q_{o1} & q_{o2} & & q_{oN} \\ q_{1o} & -q_1 & q_{12} & & q_{1N} \\ q_{20} & q_{21} & -q_2 & & q_{2N} \\ & & & & \\ q_{NO} & q_{N1} & q_{N2} & & -q_N \end{bmatrix} \qquad 2.7$$

and I is the identity matrix.

Kolmogorov equations (2.4)' and (2.5)', under initial conditions (2.6)' , have the same solution given by,

$$P(t) = e^{Qt} \qquad 2.8$$

The matrix function e^{Qt} can be evaluated using method of constituent matrices or Caley-Hamilton theorem, or may be evaluated by the convergent exponential series,

$$e^{Qt} = \sum_{k=o}^{\infty} Q^k \frac{t^k}{k!}, \quad Q^o = I \qquad 2.9$$

Kolmogorov equations, described in the above discussion, are the heart of the exponential stochastic model. It can be clearly seen that the model is completely defined by the transition rate matrix Q. The matrix Q, of course, depends on the physical stochastic system under consideration. As it has been described by (2.1) and (2.2), this matrix represents the infinitismal transition conditions of the system. Different stochastic systems that are represented by the above model are given in the next section. The emphasis will be placed there on the birth and death model due to its wide importance and relevance in many system simulation models.

III. APPLICATIONS

It has been seen in the last section that the exponential Markov model of any physical stochastic system is completely described by the transition rate matrix Q. In most applications, this matrix has simple structure and it turns out to be a band matrix. Different applications to such a model are given in references (1,3). One of the widest applications is the birth and death process.

The birth and death process itself is representative to many models such as queuing,immigration, failure of machines, and many different applications in engineering and applied science. The transition rate matrix entries for the general birth and death process are given by,

$$q_{ij} = \begin{cases} \lambda_i & , j = i + 1 \quad \text{(birth)} \\ \mu_i & , j = i - 1 \quad \text{(death)} \\ \lambda_i + \mu_i & , j = i \neq 0 \quad \text{(no change)} \\ 1 & , i = j = 0 \quad \text{(0-state is absorbing)} \end{cases} \qquad 3.1$$

where λ_i and μ_i are the i-th birth and death rates, respectively.

If $\mu_i=0$, then the process is pure birth process and if besides, $\mu_i=0$, $\lambda_i = i\lambda$ where λ is constant, then it is Yule Process. But if $\lambda_i = \lambda$ = constant and $\mu_i = 0$, the process will turn out to be the familiar Poisson process. In the case where $\lambda_i = \lambda$ and $\mu_i = \mu$, where λ and μ are appropriate constant rates, then the birth and death process yields the single-server exponential queuing system. The process will yield s-server exponential queuing system if

$$\lambda_i = \lambda$$

$$\mu_i = \begin{cases} i\mu & 1 < i < s \\ s\mu & i > s \end{cases} \qquad 3.2$$

where s is the number of servers in the system. The s-server queuing model becomes an infinite server model when $\mu_i = i\mu$ for all $i > 1$. In the case when $\lambda_i = i\lambda$ and $\mu_i = i\mu$, the general birth and death model will clearly become the so-called linear birth and death process. Besides, if $\lambda i = i\lambda + \alpha$, where α is an exponential rate of increase from external source such as immigration, the process is called linear birth and death process with immigration α.

Another model which is represented by Kolmogorov equations and which has important relevance to operational efficiency of complex engineering systems is the system reliability model. The forward equation for the replacement of failing components of unserviced system is given by [1]

$$P'_k(t) = -h(t)\, P_k(t) + h(t)\, P_{k-1}(t), \quad k>0$$

$$P'_o(t) = -h(t)\, P_o(t) \qquad 3.3$$

where $h(t)$ is component hazard rate and $P_k(t)$ is the probability of replacement k components in the interval $(0,t)$. If the system has N components and the failure distribution of the k-th component is negative exponential with hazard rate $\lambda\mu$ and service rate μ, then the reliability model is represented by birth and death process where the state of the system is the number of failed components. If we assume that at least ℓ components should be working for the system to be operative then the forward Kolmogorov's equations are given by,

$$P'_k(t) = -(\lambda_n+\mu_n)\, P_k(t) = \lambda_{n-1}\, P_{n-1}(t) + \mu_{k+1} P_{k+1}(t),$$

$$0<k\leq N-\ell \qquad 3.4$$

$$P'_{N-\ell+1}(t) = -\mu_{N-\ell+1} P_{N-\ell+1}(t) + \lambda_{N-\ell}\, P_{N-\ell}(t)$$

$$P'_o(t) = -\lambda_o P_o(t) + \mu_1 P_1(t)$$

$$P_o(0) = 1$$

It may be appropriate to mention here some of the other applications which fit the above exponential Markov model which occur in some engineering and industrial applications. Some examples are, electron emission from the cathod to the anode in electron tubes, nuclear growth, telephone traffic problems, traffic flow and pedestrian traffic problems, airport simulations of plane arrivals and departures and many, many others.

IV. STEADY STATE TRANSITION

The Markov exponential model was characterized by Kolmogorov equations (2.4), (2.5) and (2.6) or (2.4)', @2.5)', and (2.6)' in matrix form. It is necessary to solve these equations to obtain the transition probabilities $P_{ij}(t)$. This, in most cases, is very difficult and especially in the case when the number of the system states is large. In many applications, we are not interested in the dynamic transient behavior of the system, but rather in its limiting steady state behaviors as $t \to \infty$. In these applications, we assume under condition of process irreducibility (1) that the limit

$$\lim_{t\to\infty} P_{ij}(t) = P_j \qquad 4.1$$

exists and independent of the starting state i. If this is the case, then $P'_{ij}(t)$ converges to zero and the forward Kolmogorov equations (2.4) become

$$o = -q_j P_j + \sum_{k\neq j} q_{kj} P_k, \quad j = 1,2,\ldots N \qquad 4.2$$

and in the matrix form

$$PQ = 0 \qquad 4.3$$

where P is the row vector $(P_1, P_2, \ldots, P_N)$.

Equations (4.2) which may be written in the form,

$$q_j\, P_j = \sum_{k\neq j} q_{kj}\, P_k, \quad j = 1,2,\ldots \qquad 4.4$$

are called the balance equations [18] since they equate or balance the rate at which the process enters state j with the rate at which it leaves this state. The balance equations with

$$\sum_{j=0} P_j = 1 \qquad 4.5$$

will yield the steady state probability vector P. The j-th entry P_j of P, denotes the limiting probability that the system is in state j. The limiting probability vector is all that we need to know about the system. Its knowledge will provide us with the mean state of the system, the mean time time that the system spends in state j, stability of the model and all other statistics that may describe the steady state behavior of the system. For example, the balance equations for the general birth and death process can be easily obtained as, [18]

$$(\lambda_j + \mu_j)\, P_j = \mu_{j+1}\, P_{j+1} + \lambda_{j-1} P_{j-1} \quad ij> 1.$$

$$\lambda_o P_o = \mu_1 P_1 \qquad 4.6$$

Equations (4.5) and (4.6) will yield directly to the probability P_j; i.e.

$$P_j = \frac{\lambda_o\, \lambda_1 \ldots \lambda_{j-1}}{\mu_1\, \mu_2 \ldots \mu_j \left(1 + \sum_{j=1}^{\infty} \frac{\lambda_o\, \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j}\right)}, \quad j>1, \qquad 4.7$$

for $N = \infty$. It can be easily seen, from (4.7) that the limiting probabilities P_j exist if

$$\sum_{j=1}^{\infty} \frac{\lambda_o\, \lambda_1 \ldots \lambda_{j-1}}{\mu_1\, \mu_2 \ldots \mu_j} < \infty \qquad 4.8$$

The stability condition (4.8), for single server queue becomes

$$\sum_{j=o}^{\infty} \left(\frac{\lambda}{\mu}\right)^j < \infty \qquad 4.9$$

or equivalently,

$$\frac{\lambda}{\mu} < 1 \qquad 4.10$$

For the s-server queuing system condition (4.8) reduces to

$$\sum_{j=1}^{\infty} \frac{\lambda^j}{j!\, \mu^j} + \sum_{j=s+1}^{\infty} \frac{\lambda^j}{(s\mu)^j} < \infty \qquad 4.11$$

or equivalently,

$$\frac{\lambda}{s\mu} < 1 \qquad 4.12$$

Similar results may be obtained for examples and models treated in section three.

V. CONCLUSION

It has been shown that many stochastic system models may be completely represented, under some mild restrictions, by discrete-state Markov model which was described by Kolmogorov forward and backward equations. The model is very flexible and in the steady state.

It is reduced to the balance equations which yield an illustrative analytic solution for some of the special cases treated. In the general case, analytic solutions cannot be easily obtained and even the transition rates may not be available and therefore simulation will be necessary to obtain adequate results.

REFERENCES

(1) Baychi, T. P. and J. G. C. Templeton, Numerical Methods in Markov Chains and Bulk Queues, Springer-Verlag, New York, 1972.

(2) Bhat, U. Narayan, Elements of Applied Stochastic Processes, John Wiley, New York, 1972.

(3) Breiman, Leo, Probability and Stochastic Processes: with a View Toward Application, Houghton Mifflin Company, Boston, 1969.

(4) Baracha-Reid, A.T., Elements of the Theory of Markov Processes and Their Applications, McGraw-Hill, New York, 1960.

(5) Cox, D.R. and Miller, H.D., The Theory of Stochastic Processes, Wiley, New York, 1965.

(6) Doob, J, Stochastic Processes, Wiley, New York, 1953.

(7) Feller, William, An Introduction to Probability Theory and Its Applications , Volume One, Wiley, New York, 1968.

(8) Ibid.____________________, Volume 2, Wiley, New York, 1971.

(9) Frame, J.S. "Matric Functions and Applications," Part IV, IEEE Spectrum, June, 1964.

(10) IEEE Tutorial Course: "Probability Analysis of Power Systems Reliability", IEEE, New York, 1971.

(11) Karlin, S. , A First Course in Stochastic Processes, Academic Press, New York, 1966.

(12) Kiviat, P. J., "Systems Simulation", International Symposium on System Engineering and Analysis, Volume I, Purdue University, 1972.

(13) Nayler, T.H. (ed.) The Design of Computer Simulations Experiments, Duke University Press, Durham, North Carolina, 1969.

(14) Nuamaker, J.F., Jr., Pomeranz, J.E., "Principles of Discrete System Simulation", International Symposium on System Engineering and Analysis, Volume I, Purdue University, 1972.

(15) Papoulis, Athanasios, Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York, 1965.

(16) Porzon, E.,Stochastic Processes, Holden-Day, San Francisco, 1967.

(17) Pala, R.E., et al., Waiting-Line Models, Reinhold Publishing Corporation, New York, 1967.

(18) Ross, Sheldon M., Applied Probability Models with Optimization Applications, Holden-Day, San Francisco, 1970.