Diffusion Approximation as a Modelling Tool

Tadeusz Czachórski[1] and Ferhan Pekergin[2]

[1] Institute of Theoretical and Applied Informatics
Polish Academy of Sciences
Baltycka 5, 44–100 Gliwice, Poland
tadek@iitis.gliwice.pl
[2] LIPN, Université Paris-Nord, 93430 Villetaneuse, France
pekergin@lipn.univ-paris13.fr

Abstract. Diffusion theory is already a vast domain of knowledge. This tutorial lecture does not cover all results; it presents in a coherent way an approach we have adopted and used in analysis of a series of models concerning evoluation of some traffic control mechanisms in computer, especially ATM, networks. Diffusion approximation is presented from engineer's point of view, stressing its utility and commenting numerical problems of its implementation. Diffusion approximation is a method to model the behavior of a single queueing station or a network of stations. It allows one to include in the model general sevice times, general (also correlated) input streams and to investigate transient states, which, in presence of bursty streams (e.g. of multimedia transfers) in modern networks, are of interest.

Keywords: diffusion approximation, queueing models, performance evaluation, transient analysis.

1 Single G/G/1 Station

1.1 Preliminaries

Let $A(x)$, $B(x)$ denote the interarrival and service time distributions at a service station. The distributions are general but not specified, the method requires only their two first moments. The means are: $E[A] = 1/\lambda$, $E[B] = 1/\mu$ and variances are $\mathrm{Var}[A] = \sigma_A^2$, $\mathrm{Var}[B] = \sigma_B^2$. Denote also squared coefficients of variation $C_A^2 = \sigma_A^2\lambda^2$, $C_B^2 = \sigma_B^2\mu^2$. $N(t)$ represents the number of customers present in the system at time t.

Define

$$\tau_k = \sum_{i=1}^{K} a_i,$$

where a_i are time intervals between arrivals. We assume that they are independent and indentically distributed random variables, hence, according to the central limit theorem, distribution of a variable

$$\frac{T_k - k\,\lambda}{\sigma_A\sqrt{k}}$$

D. Kouvatsos (Ed.): Next Generation Internet, LNCS 5233, pp. 447–476, 2011.

tends to standard normal distribution as $k \to \infty$:

$$\lim_{k\to\infty} P[\frac{T_k - k\,\lambda}{\sigma_A \sqrt{k}} \le x] = \Phi(x), \quad \text{where} \quad \Phi(x) = \frac{1}{\sqrt{2\Pi}} \int_{-\infty}^{x} e^{-\frac{\xi^2}{2}}\, d\xi.$$

hence for a large k: $P[\tau_k \le x\sigma_A\sqrt{k} + k/\lambda] \approx \Phi(x)$. Denote $t = x\sigma_A\sqrt{k} + k/\lambda$, or $k = t\lambda - x\sigma_A\sqrt{k}\lambda$ and for large values of k, $k \approx t\lambda$ or $\sqrt{k} \approx \sqrt{t\lambda}$. Denote by $H(t)$ the number of customers arriving to the station during a time t; note that $P[H(t) \ge k] = P[\tau_k \le t]$, hence

$$\Phi(x) \approx P[\tau_k \le x\sigma_A\sqrt{k} + k/\lambda] = P[H(t) \ge k] =$$
$$= P[H(t) \ge t\lambda - x\sigma_A\sqrt{t\lambda}\,\lambda] = P\left[\frac{H(t) - t\lambda}{\sigma_A\sqrt{t\lambda}\,\lambda} \ge -x\right]$$

As for the normal distribution $\Phi(x) = 1 - \Phi(-x)$, then $P[\xi \ge -x] = P[\xi \le x]$, and

$$P\left[\frac{H(t) - t\lambda}{\sigma_A\sqrt{t\lambda}\,\lambda} \le x\right] \approx \Phi(x)\,,$$

that means that the number of customers arriving at the interval of length t (sufficiently long to assure large k) may be approximated by the normal distribution with mean λt and variance $\sigma_A^2\lambda^3 t$. Similarly, the number of customers served in this time is approximately normally distributed with mean μt and variance $\sigma_B^2\mu^3 t$, provided that the server is busy all the time. Consequently, the changes of $N(t)$ within interval $[0, t]$, $N(t) - N(0)$, have approximately normal distribution with mean $(\lambda - \mu)t$ and variance $(\sigma_A^2\lambda^3 + \sigma_B^2\mu^3)t$.

Diffusion approximation [54,55] replaces the process $N(t)$ by a continuous diffusion process $X(t)$ the incremental changes of wich $dX(t) = X(t + dt) - X(t)$ are normally distributed with the mean βdt and variance αdt, where β, α are coefficients of the diffusion equation

$$\frac{\partial f(x,t;x_0)}{\partial t} = \frac{\alpha}{2}\frac{\partial^2 f(x,t;x_0)}{\partial x^2} - \beta\frac{\partial f(x,t;x_0)}{\partial x} \tag{1}$$

which defines the conditional pdf $f(x,t;x_0) = P[x \le X(t) < x + dx \mid X(0) = x_0]$ of $X(t)$. The both processes $X(t)$ and $N(t)$ have normally distributed changes; the choice $\beta = \lambda - \mu$, $\alpha = \sigma_A^2\lambda^3 + \sigma_B^2\mu^3 = C_A^2\lambda + C_B^2\mu$ ensures the same ratio of time-growth of mean and variance of these distributions.

More formal justification of diffusion approximation is in limit theorems for $G/G/1$ system given by Iglehart and Whitte [42,43]. If $\hat{N}_n$ is a series of random variables deriven from $N(t)$:

$$\hat{N}_n = \frac{N(nt) - (\lambda - \mu)nt}{(\sigma_A^2\lambda^3 + \sigma_B^2\mu^3)\sqrt{n}}\,,$$

then the series is weakly convergent (in the sense of distribution) to ξ where $\xi(t)$ is a standard Wiener process (i.e. diffusion process with $\beta = 0$ i $\alpha = 1$) provided that $\varrho > 1$, that means if the system is overloaded and has no equilibrium state. In the case