KALMAN FILTERING OF MIXTURE SYSTEMS WITH INTERVAL DYNAMICS

Alexandru Murgu
Department of Mathematical Information Technology
University of Jyväskylä
FIN-40351 Jyväskylä, FINLAND

Abstract

In this paper an interval dynamical modelling for Kalman filtering of the mixture models associated to the multistream traffic aggregation in communication links in general flow networks is considered. The purpose is to build a robust forecasting device working under uncertainty of the traffic classes assignment. If some system parameters such as certain elements of the system matrix are not precisely known or gradually change with time, then the classical Kalman filtering algorithm cannot be applied directly. In this case, a robust version of the Kalman filtering that has the ability of handling the traffic class uncertainty is needed. The interval data vector obtained in the interval Kalman filtering is an uncertain interval vector before the data are actually being obtained. It will be an ordinary constant vector after it has been actually measured and realized.

Keywords: Interval Dynamics, Kalman Filtering, Mixture Models, Covariance Matrix, Traffic Patterns.

1. Introduction

The multiplexing of various traffic stream allows the minimization of the transmission costs and time delays is considering the decomposition of the traffic patterns based on hierarchical mean clustering of the traffic component rates. The multiclass service systems are becoming even more popular with the penetration of mobile internet applications. The competition to ensure a highly reliable and guaranteed quality-of-service parameters is forcing a new perspective in designing the real-life network operation regimes where the uncertainty of the current traffic classes assignment should be tolerated ([2]). The interval modelling is solution of these problems and it is motivated by flexibility in implementing adaptive flow multiplexing strategies that are able to track the relevant statistics of the real flow network while meeting the uncertain traffic demand. The ATM networks have been designed by the service provider industry (telcos) offers state-of-the-art quality of service (QoS)., where the ATM connections are described by the bit rates or bandwidth (constant, variable, unspecified and available bit rates). The ATM traditional strenghts are that it can cope well with delay sensitive applications and provide varying classes of service (allowing the service providers to offer a suite of service level agreements). This is achieved by ATM transforming the "connectionless" nature of the IP traffic into a connection-oriented mode. The IP is connectionless, that is, it operates on a "best-effort" principle where packets are routed individually using the destination IP adress inside every packet. The routes to be followed by these packets are determined by IP routing protocols such as open shortes path firs (OSPF). This directs packets onto the shortest path and makes no distinction between voice and data. This can lead to hyper aggregation on one path as OSPF does not take into account how much other traffic is traversing one particular link at any one time. In the ATM environment, the IP traffic is directed over predetermined routes and so it avoids the downfalls of best effort.

2. Mixture Systems for Multistream Aggregation

Let us consider the multistream traffic patterns represented as classes of stochastic processes, $C_k,\ k = 1,\ 2,\dots,K$. Assume that we have a multiple hypothesis test, each hypothesis corresponding to one class (for which specific control strategies are designed), that is

$$H_k:\ \mathbf{x} \in C_k,\ k = 1,\ 2,\dots,K \tag{1}$$

The problem associated to the traffic stream discrimination in the aggregation set ([4]) is to divide the pattern space Ω_x into K disjoint decision regions, that is, $\Omega_1,\ \Omega_2,\dots,\ \Omega_K$. Let p_k be the prior probability that $\mathbf{x}$ belongs to C_k and $f_k(\mathbf{x}) = f(\mathbf{x}|C_k)$ is the conditional density function or the likelihood function. We assume that the costs of an error of different kinds are identical and set the cost $c = 1$. Then the risk associated with an arbitary decision function $\delta'(\mathbf{x})$ and decision regions, $\Omega'_1,\ \Omega'_2,\dots,\ \Omega'_K$, is given by

$$\rho(\delta') = \sum_{k=1}^{K} \int_{\Omega'_k} \sum_{l \neq k} p_l f_l(\mathbf{x}) d\mathbf{x} \tag{2}$$

and represents the error probability. The Bayes classifier ([8]) that minimizes ρ will divide Ω_x into K disjoint regions, with each region Ω_j consisting of values of $\mathbf{x}$ such that

$$p_j f_j(\mathbf{x}) \geq p_k f_k(\mathbf{x}),\ \text{for all}\ k \neq j \tag{3}$$

When a pattern vector $\mathbf{x}$ is observed, we first calculate the likelihood function $f_k(\mathbf{x})$, multiply it by the prior probability p_k, select the one with the largest value and decide that $\mathbf{x}$ belongs to the class corresponding to the maximum $p_k f_k(\mathbf{x})$. Using the Bayesian rule, we have

$$p_k f_k(\mathbf{x}) = f(\mathbf{x}|C_k)P(C_k) = P(C_k|\mathbf{x})f(\mathbf{x}) \tag{4}$$

where

$$f(\mathbf{x}) = \sum_{k=1}^{K} p_k f_k(\mathbf{x}) \tag{5}$$

is the probability density of $\mathbf{x}$ (mixture of density function). The comparison of $p_k f_k(\mathbf{x})$ for different k is equivalent to comparing the posterior probability $P(C_k|\mathbf{x})$. For equal cost, the Bayesian classifier is essentially a maximum posterior probability classifier. In the case of multiple classes of observations, assume a sequence of pattern vectors $\mathbf{x}_1,\ \mathbf{x}_2,\dots,\mathbf{x}_n$ is observed and that all vectors in the sequence come from the same class, either C_1 or C_2. The two hypotheses are

$$H_1:\ \mathbf{x}_i \in C_1,\ i = 1,\ 2,\dots,n \tag{6}$$

$$H_2:\ \mathbf{x}_i \in C_2,\ i = 1,\ 2,\dots,n \tag{7}$$

Using the likelihood ratio test

$$\Lambda_n = \Lambda(\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_n) = \frac{f_2(\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_n)}{f_1(\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_n)} \tag{8}$$

where $f_k(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ is the likelihood function of the sequence given that $\mathbf{x}_i \in C_k$, $i = 1,\ 2, \dots, n$, since the samples are observed randomly, it is reasonable to assume that $\mathbf{x}_1,\ \mathbf{x}_2, \dots, \mathbf{x}_n$ are independent and identically distributed, that is,

$$f_k(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \prod_{i=1}^{n} f_k(\mathbf{x}_i), \quad k = 1,\ 2 \tag{9}$$

The members of the finite sequence $\mathbf{x}_1,\ \mathbf{x}_2, \dots, \mathbf{x}_n$ belong to the same class, either C_1 or C_2. When they belong to C_1, the random samples are all distributed according to $f(\mathbf{x}_i | C_1) = f_1(\mathbf{x}_i)$ and the independence should be in terms of the conditional densities with $k = 1$. Similar for C_2. This type of independence is called *conditional independence*. Let us assume that each $f_k(\mathbf{x})$ is an N-dimensional Gaussian density function with mean vector $\mathbf{m}_k$ and covariance matrix $\mathbf{R}_k$, that is,

$$f_k(\mathbf{x}) = G(\mathbf{x}; \mathbf{m}_k, \mathbf{R}_k) = (2\pi)^{-N/2} |\mathbf{R}_k|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{m}_k)^T \mathbf{R}_k^{-1} (\mathbf{x} - \mathbf{m}_k)\right] \tag{10}$$

For two classes with $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}$ and $p_1 = p_2 = 0.5$, the discriminant function is a linear function, that is,

$$-D(\mathbf{x}) = \mathbf{x}^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) + \frac{1}{2}(\mathbf{m}_2^T \mathbf{R}^{-1} \mathbf{m}_2 - \mathbf{m}_1^T \mathbf{R}^{-1} \mathbf{m}_1) \tag{11}$$

and the decision surface is a hyperplane. If we define the statistic

$$s = \mathbf{x}^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \tag{12}$$

and let

$$s_0 = \frac{1}{2}(\mathbf{m}_2^T \mathbf{R}^{-1} \mathbf{m}_2 - \mathbf{m}_1^T \mathbf{R}^{-1} \mathbf{m}_1) \tag{13}$$

the classification will be based on whether $s < s_0$. Given that $\mathbf{x} \in C_1$, we have the following first and second order conditional expectations

$$E\left[s | C_1\right] = \int_{\Omega_x} \mathbf{x}^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) f_1(\mathbf{x}) d\mathbf{x} = \mathbf{m}_1^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \tag{14}$$

$$\begin{aligned} E\left[s^2 | C_1\right] &= \int_{\Omega_x} (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{R}^{-1} \mathbf{x}\mathbf{x}^T (\mathbf{m}_2 - \mathbf{m}_1) f_1(\mathbf{x}) d\mathbf{x} = \\ &= (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{R}^{-1} (\mathbf{R} + \mathbf{m}_1 \mathbf{m}_1^T) \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \end{aligned} \tag{15}$$

The variance is

$$\mathrm{Var}\left[s | C_1\right] = E\left[s^2 | C_1\right] - \left(E\left[s | C_1\right]\right)^2 = (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \tag{16}$$

Similarly,

$$E\left[s | C_2\right] = \mathbf{m}_2^T \mathbf{R}^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \tag{17}$$

$$\mathrm{Var}\left[s | C_2\right] = \mathrm{Var}\left[s | C_1\right] \tag{18}$$

We can define a measure of separation

$$d^2 = \frac{\left(E\left[s\middle|C_2\right] - E\left[s\middle|C_1\right]\right)^2}{\mathrm{Var}\left[s\middle|C_1\right]} = (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{R}^{-1}(\mathbf{m}_2 - \mathbf{m}_1) \tag{19}$$

which is called the "signal-to-noise" ratio. Since s is a linear combination of the components of the Gaussian random vector $\mathbf{x}$, it is Gaussian distributed with the density function

$$h_1(s) = h(s\middle|C_1) = G(s; \mathbf{m}_1^T \mathbf{R}^{-1}(\mathbf{m}_2 - \mathbf{m}_1), d^2) \tag{20}$$

3. Systems with Interval Dynamics

3.1. General Framework

In the discrimination model of the multistream aggregation, some system parameters such as certain elements of the system's covariance matrix are not precisely known or gradually change with time, then the Kalman filtering algorithm cannot be applied directly. We consider a version of robust Kalman filtering scheme ([1]). Let us consider the nominal system

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k \tag{21}$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \tag{22}$$

where $\mathbf{A}_k$, $\mathbf{B}_k$ and $\mathbf{C}_k$ are known $n \times n$, $n \times p$ and $q \times n$ matrices respectively, describing the traffic stream clasees dynamics, with $1 \le p,\ q \le n$. We assume that

$$E\left[\mathbf{u}_k\right] = \mathbf{0},\ \ E\left[\mathbf{u}_k \mathbf{u}_k^T\right] = \mathbf{Q}_k \delta_{kl} \tag{23}$$

$$E\left[\mathbf{v}_k\right] = \mathbf{0},\ \ E\left[\mathbf{v}_k \mathbf{v}_k^T\right] = \mathbf{R}_k \delta_{kl} \tag{24}$$

$$E\left[\mathbf{u}_k \mathbf{v}_l^T\right] = \mathbf{0},\ E\left[\mathbf{x}_0 \mathbf{u}_k^T\right] = \mathbf{0},\ E\left[\mathbf{x}_0 \mathbf{v}_k^T\right] = \mathbf{0} \tag{25}$$

for all $k,\ l = 0,\ 1, \ldots$, with $\mathbf{Q}_k$ and $\mathbf{R}_k$ being positive definite and symmetric matrices. If all the constant matrices $\mathbf{A}_k$, $\mathbf{B}_k$ and $\mathbf{C}_k$ are known, then the Kalman filter can be applied to the nominal system, which yields the optimal estimates $\left\{\hat{\mathbf{x}}_k\right\}$ using the measurement data $\left\{\mathbf{y}_k\right\}$ in a recursive scheme. If some of the elements of this system are unknown or uncertain, a modification of the entire setting for filtering is necessary. Suppose that the uncertain parameters are only known to be bounded, that is,

$$\mathbf{A}_k^I = \mathbf{A}_k + \Delta\mathbf{A}_k = \left[\mathbf{A}_k - \left|\Delta\mathbf{A}_k\right|,\ \mathbf{A}_k + \left|\Delta\mathbf{A}_k\right|\right] \tag{26}$$

$$\mathbf{B}_k^I = \mathbf{B}_k + \Delta\mathbf{B}_k = \left[\mathbf{B}_k - \left|\Delta\mathbf{B}_k\right|,\ \mathbf{B}_k + \left|\Delta\mathbf{B}_k\right|\right] \tag{27}$$

$$\mathbf{C}_k^I = \mathbf{C}_k + \Delta\mathbf{C}_k = \left[\mathbf{C}_k - \left|\Delta\mathbf{C}_k\right|,\ \mathbf{C}_k + \left|\Delta\mathbf{C}_k\right|\right] \tag{28}$$

for $k = 0,\ 1, \ldots$, where $\left|\Delta\mathbf{A}_k\right|$, $\left|\Delta\mathbf{B}_k\right|$ and $\left|\Delta\mathbf{C}_k\right|$ are constant bounds for the unknowns. The corresponding system

$$\mathbf{x}_{k+1} = \mathbf{A}_k^I \mathbf{x}_k + \mathbf{B}_k^I \mathbf{u}_k \tag{29}$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \tag{30}$$

is called the interval system. The physical realization of the interval matrices (26)-(28) are prescribing the traffic stream class assignment during the real-life operation of the network.

Proposition 1. Let $I_1,\ I_2,\ J_1$ and J_2 be intervals with

$$I_1 \subset J_1 \text{ and } I_2 \subset J_2 \tag{31}$$

Then for any operation $* \in \{+,-,\cdot,/\}$, it follows that

$$I_1 * I_2 \subset J_1 * J_2 \tag{32}$$

∎

Corollary 1. (Monotonic Inclusion Property). Let I and J be intervals and let $x \in I$ and $y \in J$. Then

$$x * y \subseteq I * J \tag{33}$$

for all $* \in \{+,-,\cdot,/\}$. ∎

3.2. Rational Interval Functions

Let S_1 and S_2 be intervals in R and $f: S_1 \to S_2$ be a real-valued function. Denote by Σ_{S_1} and Σ_{S_2} the families of all subintervals of S_1 and S_2, respectively. The interval-to-interval function, $f^I : \Sigma_{S_1} \to \Sigma_{S_2}$ is defined by

$$f^I(X) = \left\{ f(x) \in S_2 \middle| x \in X,\ X \in \Sigma_{S_1} \right\} \tag{34}$$

is called the *united extension* of the point-to-point function f on S_1. Its range is

$$f^I(X) = \bigcup_{x \in X} \{f(x)\} \tag{35}$$

Proposition 2. The following property of the united extension $f^I : \Sigma_{S_1} \to \Sigma_{S_2}$ is a direct consequence of the above definition

$$X,\ Y \in \Sigma_{S_1} \text{ and } X \subseteq Y \ \Rightarrow\ f^I(X) \subseteq f^I(Y) \tag{36}$$

∎

An interval-to-interval function F of n variables, $X_1,\ X_2, \ldots, X_n$, is said to have the *monotonic inclusion property* if

$$X_i \subseteq Y_i,\ i = 1, \ldots, n \ \Rightarrow\ F(X_1, \ldots, X_n) \subseteq F(Y_1, \ldots, Y_n) \tag{37}$$

Not all the interval-to-interval functions have this property. All united extensions have the monotonic inclusion property. An interval-to-interval function is called interval function for simplicity. Interval vectors} and interval matrices are similarly defined. An interval function is said to be *rational* and it is called *a rational interval function*, if its values are defined by a finite sequence of intervals arithmentic operations. All the rational interval functions have the monotonic inclusion property. Let $f = f(x_1,...,x_n)$ be an ordinary n-variable real-valued function and be intervals. An interval function, $F = F(X_1, X_2,..., X_n)$, is said to be an *interval extension* of f if

$$F(x_1,...,x_n) = f(x_1, x_2,..., x_n) \tag{38}$$

for all $x_i \in X_i$, $i = 1,...,n$. Not all the interval extensions have the monotonic inclusion property. ∎

Proposition 3. If F is an interval extension of f with the monotonic inclusion property, then the united extension f^I of f satisfies

$$f^I(X_1,...,X_n) \subseteq F(X_1,...,X_n) \tag{39}$$

∎

Corollary 2. If F is a rational interval function and is an interval extension of f, then

$$f^I(X_1,...,X_n) \subseteq F(X_1,...,X_n) \tag{40}$$

∎

Corollary 2 provides a means of finite evaluation of upper and lower bounds on the value-range of an ordinary rational function over an n-dimensional rectangular domain in R^n ([3]).

3.3. Interval Expectation and Variance

Let $f(x)$ be an ordinary function defined on an interval X. If f satisfies the ordinary Lipschitz condition

$$\left|f(x) - f(y)\right| \le L|x - y| \tag{41}$$

for some positive constant L which is independent on $x,\ y \in X$, then the united extension f^I of f is said to be a *Lipschitz interval extension* of f over X. Let N be a positive integer and let us subdivide an interval $[a,\ b] \subseteq X$ into N subintervals, $X_1 = \left[\underline{X}_1,\ \overline{X}_1\right]$, …, $X_N = \left[\underline{X}_N,\ \overline{X}_N\right]$, such that

$$a = \underline{X}_1 < \overline{X}_1 = \underline{X}_2 < \overline{X}_2 = \cdots = \underline{X}_N < \overline{X}_N = b \tag{42}$$

For any $f \in B(X)$, let F be a Lipschitz interval extension of f defined on all X_i, $i = 1,...,N$. Assume that F satisfies the monotonic inclusion property. Using the notation

$$S_N(F;[a,\ b]) = \frac{b-a}{N}\sum_{i=1}^{N} F(X_i) \tag{43}$$

we have

$$\int_a^b f(t)dt = \bigcap_{N=1}^{\infty} S_N(F;[a,\ b]) = \lim_{N\to\infty} S_N(F;[a,\ b]) \tag{44}$$

If we recursively define

$$Y_{k+1} = S_{k+1} \cap Y_k, k = 1,\ 2, \ldots \tag{45}$$

where $Y_1 = S_1$ and $S_k = S_k(F;[a,\ b])$, then $\{Y_k\}$ is a nested sequence of intervals that conveges to the exact value of the integral in the equation (44). Let X be an interval of real-value random variables of interest and let

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right], \quad x \in X \tag{46}$$

be an ordinary Gaussian density function with known mean μ_x and variance $\sigma_x > 0$. Then $f(x)$ has a Lipschitz interval extension. The *interval expectation* is defined as

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right] dx,\ x \in X \tag{47}$$

and the *interval variance* is defined as

$$\begin{aligned} \mathrm{Var}[X] = E\left[\left(X - E[X]\right)^2\right] = \int_{-\infty}^{\infty} (x-\mu_x)^2 f(x)dx = \\ = \int_{-\infty}^{\infty} \frac{(x-\mu_x)^2}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right] dx,\ x \in X \end{aligned} \tag{48}$$

The integrals in (47) and (48) are both well defined. This can be easily verified based on the definite integral defined above, with $a \to -\infty$ and $b \to \infty$. The *conditional interval expectation* with respect to another real interval Y of real-valued random variables is defined as

$$\begin{aligned} E[X|y \in X] = \int_{-\infty}^{\infty} xf(x|y)dx = \int_{-\infty}^{\infty} x\frac{f(x,y)}{f(y)}dx = \\ = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_{xy}} \exp\left[-\frac{(x-\mu_{xy})^2}{2\sigma_{xy}^2}\right] dx, \quad x \in X \end{aligned} \tag{49}$$

and the *conditional variance*

$$\begin{aligned} \mathrm{Var}[X|y \in Y] = E\left[\left(x-\mu_x\right)^2 |y \in Y\right] = \int_{-\infty}^{\infty} (x - E[x|y \in Y])^2 f(x|y)dx = \\ = \int_{-\infty}^{\infty} (x - E[x|y \in Y])^2 \frac{f(x,y)}{f(y)}dx = \\ = \int_{-\infty}^{\infty} \frac{(x - E[x|y \in Y])^2}{\sqrt{2\pi}\tilde{\sigma}} \exp\left[-\frac{(x-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right] dx,\ x \in X \end{aligned} \tag{50}$$

Proposition 4. In the above equations, we have

$$\tilde{\mu} = \mu_x + \frac{(y-\mu_y)\sigma_{xy}^2}{\sigma_y^2} \tag{51}$$

$$\tilde{\sigma}^2 = \sigma_x^2 - \frac{\sigma_{xy}^2 \sigma_{yx}^2}{\sigma_y^2} \tag{52}$$

where

$$\sigma_{xy}^2 = \sigma_{yx}^2 = E[XY] - E[X]E[Y] = E[xy] - E[x]E[y],\ x,\ y \in X \tag{53}$$

▮

Proposition 5. The conditional expectation is given by

$$E\left[X \middle| y \in Y\right] = E[x] - (y - E[y])^2 \frac{\sigma_{xy}^2}{\sigma_y^2},\ x \in X \tag{54}$$

and the conditional variance is given by

$$\mathrm{Var}\left[X \middle| y \in Y\right] = \mathrm{Var}[x] - \frac{\sigma_{xy}^2 \sigma_{yx}^2}{\sigma_y^2},\ x \in X \tag{55}$$

▮

4. Interval Kalman Filtering for Traffic Class Estimation

The interval system previously defined has an upper boundary system defined by all upper bounds of elements of its interval matrices

$$\mathbf{x}_{k+1} = (\mathbf{A}_k + |\Delta \mathbf{A}_k|)\mathbf{x}_k + (\mathbf{B}_k + |\Delta \mathbf{B}_k|)\mathbf{u}_k \tag{56}$$

$$\mathbf{y}_k = (\mathbf{C}_k + |\Delta \mathbf{C}_k|)\mathbf{x}_k + \mathbf{v}_k \tag{57}$$

and a lower boundary system using all lower bounds of the elements of its interval matrices

$$\mathbf{x}_{k+1} = (\mathbf{A}_k - |\Delta \mathbf{A}_k|)\mathbf{x}_k + (\mathbf{B}_k - |\Delta \mathbf{B}_k|)\mathbf{u}_k \tag{58}$$

$$\mathbf{y}_k = (\mathbf{C}_k - |\Delta \mathbf{C}_k|)\mathbf{x}_k + \mathbf{v}_k \tag{59}$$

By performing the standard Kalman filtering for these two boundary systems, the resulting two filtering trajectories do not encompass all possible optimal solutions of the interval system. There is no specific relation between these two boundary trajectories and the entire family of optimal filtering solutions ([7]). The two boundary trajectories and their neighboring ones are generally intercrossing each other due to the noise perturbations ([6]).

Interval Kalman Filtering of Traffic Streams Class Assignment

1. Main Process

$$\hat{\mathbf{x}}_0^I = E\left[\mathbf{x}_0^I\right] \tag{60}$$

$$\hat{\mathbf{x}}_k^I = \mathbf{A}_{k-1}^I \hat{\mathbf{x}}_{k-1}^I + \mathbf{G}_k^I \left[\mathbf{y}_k^I - \mathbf{C}_k^I \mathbf{A}_{k-1}^I \hat{\mathbf{x}}_{k-1}^I\right],\ k = 1,\ 2,\ldots \tag{61}$$

2. Co-Process

$$\mathbf{P}_0^I = \mathrm{Var}\left[\mathbf{x}_0^I\right] \tag{62}$$

$$\mathbf{M}_{k-1}^{I} = \mathbf{A}_{k-1}^{I}\mathbf{P}_{k-1}^{I}(\mathbf{A}_{k-1}^{I})^{T} + \mathbf{B}_{k-1}^{I}\mathbf{Q}_{k-1}(\mathbf{B}_{k-1}^{I})^{T} \quad (63)$$

$$\mathbf{G}_{k}^{I} = \mathbf{M}_{k-1}^{I}(\mathbf{C}_{k}^{I})^{T}\left[\mathbf{C}_{k}^{I}\mathbf{M}_{k-1}^{I}(\mathbf{C}_{k}^{I})^{T} + \mathbf{R}_{k}\right]^{-1} \quad (64)$$

$$\mathbf{P}_{k}^{I} = \left[\mathbf{I} - \mathbf{G}_{k}^{I}\mathbf{C}_{k}^{I}\right]\mathbf{M}_{k-1}^{I}\left[\mathbf{I} - \mathbf{G}_{k}^{I}\mathbf{C}_{k}^{I}\right]^{T} + \mathbf{G}_{k}^{I}\mathbf{R}_{k}(\mathbf{G}_{k}^{I})^{T},\ k = 1, 2, \ldots \quad (65)$$

The filtering result produced by the interval Kalman filtering scheme is a sequence of interval estimates $\left\{\hat{\mathbf{x}}_{k}^{I}\right\}$ that contains all possible optimal estimates $\left\{\hat{\mathbf{x}}_{k}\right\}$ of the state vectors $\left\{\mathbf{x}_{k}\right\}$ which the interval system may generate. The filtering result produced by this interval Kalman filtering scheme is inclusive, but generally conservative in the sense that the range of interval estimates is often unnecessarily wide in order to include all possible optimal solutions. The interval data vector $\mathbf{y}_{k}^{I}$ in the interval Kalman filtering scheme is an uncertain interval vector before its realization, i.e., before the data are actually being obtained, but it will be an ordinary constant vector after it has been measured and obtained.

5. Concluding Remarks

For the service providers, the physically provisioning of the ATM links as well as the need to manage the ATM and IP layers separately, add complexity to the network. With multi-protocol label switching (MPLS), layers 2 and 3 of the network are effectively merged into one as it allows the routing of IP packets using "enhanced" aggregated functionality to be managed from one point of the network onwards. A combination of MPLS with other IP QoS protocols, such as DiffServ and RSVP can offer carrier class IP VPNs to its customers. The interval Kalman filtering performing a task of prediction and classification of the traffic stream features in the appropriate solution of these issues.

References

1. Chui, C.K. and G. Chen (1999); Kalman Filtering with Real-Time Applications, 3rd Edition; Springer-Verlag.

2. Gu, X., Sohraby, K. and D.R. Vaham (1995); Control and Performance in Packet, Circuit and ATM Networks; Kluwer Academic Piublishers.

3. Kearfott, R.B. and Z. Xing (1994); An Interval Step Control for Continuation Methods; ; SIAM Journal on Numerical Methods, Vol. 31, No. 3 (pp. 892-914).

4. McDonald, J.C. (1990); Fundamentals of Digital Switching, 2nd Edition; Plenum Press.

5. Murgu, A. and P. Lehtinen (1999); Kalman Filtering Modelling for Traffic Control in Intelligent Networks; Proceedings of the 3rd European Conference on Numerical Mathematics and Advanced Applications, ENUMATH-99 (pp. 322-332).

6. Rex, G. and J. Rohn (1998); Sufficient Conditions for Regularity and Singularity of Interval Matrices; SIAM Journal on Matric Analysis and Apllications, Vol. 20, No. 2 (pp. 437-445).

7. Ruymgaart, P.A. and T.T. Soong (1988); Mathematics of Kalman-Bucy Filtering; Springer-Verlag.

West, M. and J. Harrison (1997); Bayesian Forecasting and Dynamic Models, 2nd Edition; Springer-Verlag.