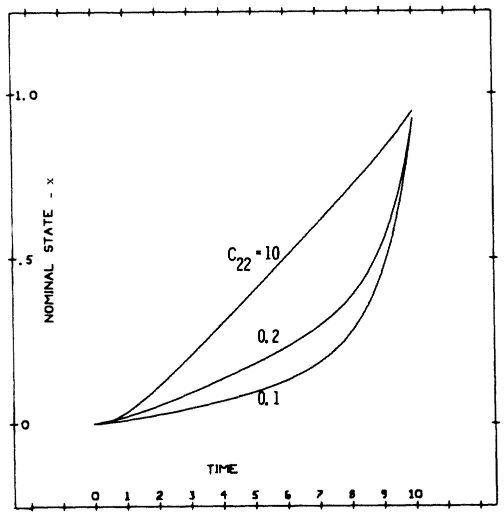

Fig. 1. Nominal state history.

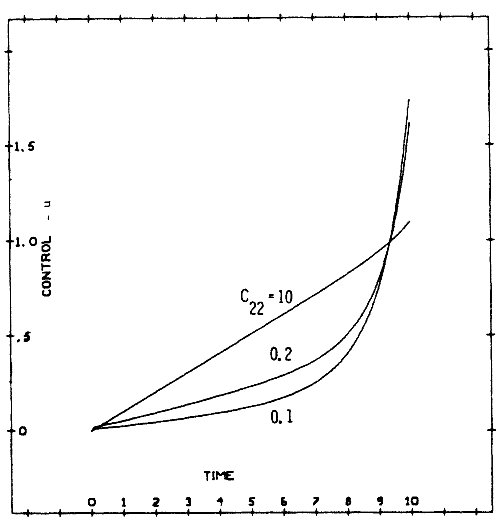

Fig. 2. Control history.

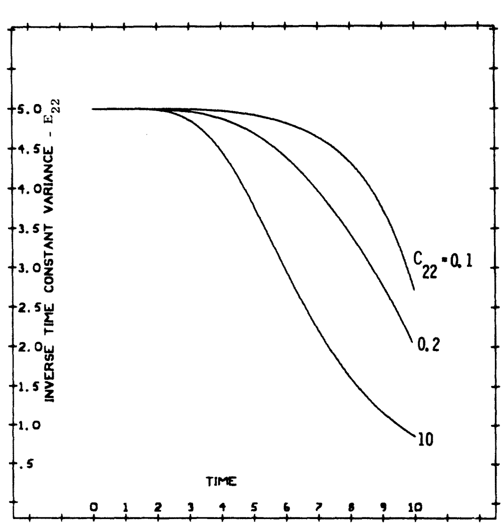

Fig. 3. State estimation performance.

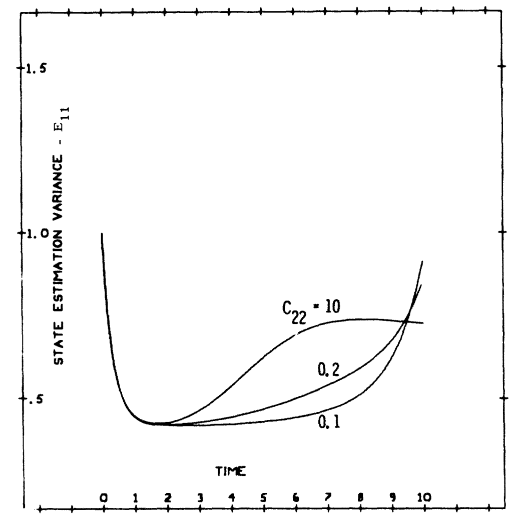

Fig. 4. Inverse time-constant performance.

Finally, this technique of shaping the open-loop trajectory has been applied to more complicated second-order systems with unknown parameters with similarly improved estimation of the states and unknown parameters at the terminal time [7].

References

[1] C. Wells, "Application of modern estimation and identification to chemical process," in *1966 Joint Automatic Control Conf., Preprints.*
[2] R. K. Mehra, "Optimal inputs for linear system identification," in *1972 Joint Automatic Control Conf., Preprints.*
[3] A. Aoki and R. Staley, "On input signal synthesis in parameter identification," *Automatica*, vol. 6, 1970.
[4] V. C. Levadi, "Design of input signal for parameter estimation," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 205–211, Apr. 1966.
[5] R. K. Mehra, "Identification of stochastic linear dynamic systems," *AIAA J.*, vol. 9, pp. 28–31, Jan. 1971.
[6] N. Nahi and D. Wallis, "Optimal inputs for parameter estimation in dynamic systems with white observation noise," in *1969 Joint Automatic Control Conf., Preprints.*
[7] G. Schmidt, "A new technique for identification and control of systems with unknown parameters," Sc.D. dissertation, M.I.T., Cambridge, Mass., Oct. 1970.
[8] G. Schmidt, "Closed-loop control of stochastic non-linear systems," *Automatica*, vol. 7, pp. 557–566, Sept. 1971.
[9] G. Schmidt, "A new technique for closed-loop control of discrete stochastic nonlinear systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-2, pp. 278–281, Apr. 1972.

An Adaptive Estimator with Learning for a Plant Containing Semi-Markov Switching Parameters

RICHARD L. MOOSE AND PAUL P. WANG

***Abstract*—An adaptive state estimator with learning has been developed to help solve the problem of state estimation of an unreliable linear system operating in Gaussian noise. By definition, the unreliable plant has certain parameters that can vary randomly within a finite set of possible values at times which are unknown to an observer. In modeling the stochastic system, it will be assumed that the variations in the plant configuration can be described by a semi-Markov process. By incorporating the semi-Markov process into a Bayesian estimation scheme an adaptive state estimator was developed which could handle the switching plant or switching environment problem without computer storage increasing as time progresses.**

Manuscript received December 30, 1971; revised December 28, 1972. This work was supported by a National Defense Fellowship. This work was presented in part at the IEEE Decision and Control Conference, Miami Beach, Fla., December 15–17, 1971.
R. L. Moose is with the Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va.
P. P. Wang is with the Department of Electrical Engineering, Duke University, Durham, N.C.

I. Introduction

At the present time, there exists a great deal of activity concerning sequential state estimation of dynamic systems. The original work in this area resulted in the familiar Kalman–Bucy filter, which represents the optimal linear approach to state estimation when the disturbances can be modeled by a Gaussian process [1], [2]. The work to be discussed herein concerns an extension of Kalman filtering theory to estimate the state of an unreliable linear dynamic plant when certain plant parameters, at times which are unknown to an observer, switch randomly within a finite set of real possible values. In addition, the random parameter variations are assumed to occur relatively slowly with respect to the response time of the system.

Briefly summarizing past efforts in this direction, we find that in 1965 Magill [3] formulated and solved the state estimation problem of a linear dynamic system that had an unknown configuration which had been randomly chosen from a set of n possible plant configurations, and once selected remained thereafter fixed. Hilborn and Lainiotis [4] next extended Magill's work from a scalar to a vector observation case and included a new method of calculating the weighting coefficients, but again it was assumed that once "learned" the system configuration would remain thereafter fixed. More recently, Ackerson and Fu [5] described an extension of the Kalman–Bucy filter which can be applied when the noise influencing the system is not Gaussian, but comes from a group of Gaussian distributions. These distributions act one at a time upon the system, with the transition from one noise source to another being described by a discrete Markov transition matrix. This interesting problem, once solved, was termed the switching environment problem. More recently, other investigators, notably, Hilborn and Lainiotis [6]–[8], and Parekh and Melsa [9], [10], have contributed toward the solution of the unreliable system problem. The resultant algorithms, however, appear difficult to implement for large-scale random varying systems, and it is felt that perhaps a different method of analysis may be worth presenting.

II. New Results

This correspondence extends the earlier work to the case where the unknown plant configuration (matrix of m state equations) randomly switches at random times between a finite set of n possible plant configurations. The rate of switching is assumed to be considerably slower than that of the observation sampling rate; also, the random switching of the unreliable plant will be modeled by a semi-Markov process. Briefly, a semi-Markov process is a probabilistic system that makes its state (plant-configuration matrix in our case) transitions according to the transition probability matrix of a conventional Markov process. However, the amount of time spent in state i before the next transmission to state j is a random variable [11]. It is this property of a random switching time that distinguishes the more general semi-Markov process from a Markov process.

By modeling the switching plant in this manner, the problem of computer storage increasing with increasing time was completely eliminated. In addition, it became possible to directly incorporate the semi-Markov switching statistics into the learning portion of an adaptive state estimator. The adaptive filter (Fig. 1) essentially consists of a bank of n Kalman filters, each matched to a possible plant configuration s_i, the outputs of which are weighted by a time-varying *a posteriori* probability. The key factor which makes this filter unique from the previous filters is that in calculating the critical weighting (*a posteriori* probabilities) coefficients a measure–predict–measure technique is used where

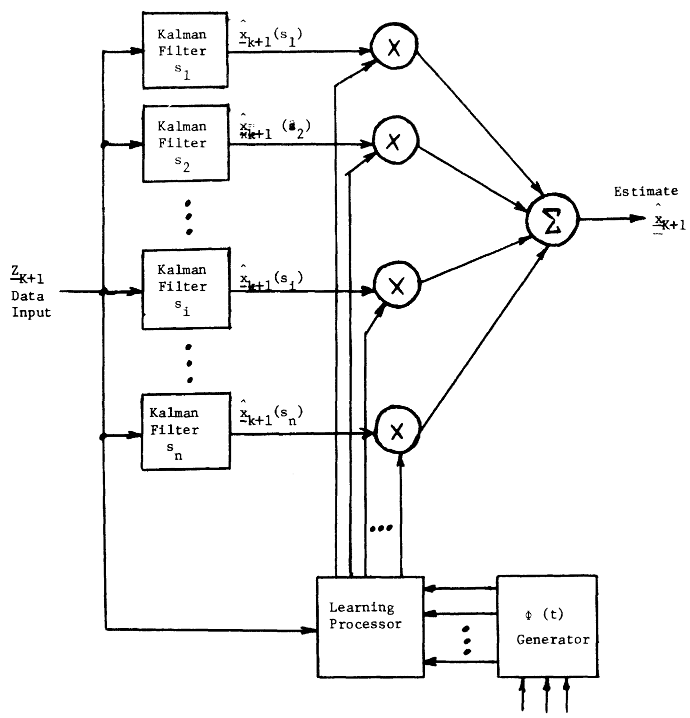

Fig. 1. Adaptive state estimator.

the semi-Markov statistics of a random starting process are used to make the intermediate predictive step.

In the verification of the adaptive filter numerous computer runs were made under many different signal-to-noise ratios and possible system configurations. The results reported on here show the operation of the adaptive filter under worst case conditions where the measurement errors are large and the plant configuration changes rapidly.

III. Derivation of Adaptive Filter

We are now in a position to develop the adaptive filter based upon the semi-Markov statistics that govern the switching plant.

The discrete form of the linear system equations can be expressed by

$$\begin{aligned} x_{k+1} &= \phi_k(s_i)x_k + \Gamma_k(s_i)w_k \\ z_{k+1} &= H_{k+1}x_{k+1} + v_{k+1} \end{aligned} \tag{1}$$

where x_{k+1} is the m state vector at time t_{k+1}; $\phi_k(s_i),\Gamma_k(s_i)$ are the state and disturbance transition matrices, both of which are functions of plant configuration s_i, $i = 1,2,\cdots,n$; w_k is an r-dimensional disturbance vector of zero-mean Gaussian inputs; v_{k+1} is an l-dimensional vector of zero-mean Gaussian measurement error; and $E[w_k v_k^T] = 0$, $E[w_k w_j^T] = Q_k\delta_{kj}$, and $E[v_k v_j^T] = R_k\delta_{kj}$.

To develop the adaptive filter a conventional Bayes estimation procedure is used. The conditional mean can be written as follows:

$$\hat{x}_{k+1} = \sum_{x_{k+1}} x_{k+1} p(x_{k+1} \mid Z_{k+1}) \tag{2}$$

where

$$p(x_{k+1} \mid Z_{k+1}) = \sum_{i=1}^{n} p(x_{k+1} \mid Z_{k+1}, s_{k+1} = s_i)p(s_{k+1} = s_i \mid Z_{k+1}). \tag{3}$$

Here Z_{k+1} represents the observation sequence from time index 1 up to time $k + 1$ and $p(s_{k+1} = s_i \mid Z_{k+1})$ is interpreted as the probability (system is in state i at time $k + 1$ | data sequence Z_{k+1}). Combining (2) and (3) and interchanging the order of summation, the following results are obtained:

$$\hat{x}_{k+1} = \sum_{i=1}^{n} \hat{x}_{k+1}(s_i)p(s_{k+1} = s_i \mid Z_{k+1}) \tag{4}$$

where

$$\hat{x}_{k+1}(s_i) = E[x_{k+1} \mid Z_{k+1}, s_{k+1} = s_i]. \tag{5}$$

Equation (5) can be closely approximated by a conventional Kalman filter due to the extremely low probability (in the order of 0.01 for many practical cases) of system switching during any one sample interval. A further area of investigation would be to study the effect on performance of the Kalman filtering assumption if the probability of transition were increased by a significant amount. Thus far there is nothing new in the derivation, and (4) will be modeled by a bank of n Kalman filters, each matched to a particular system configuration s_i and multiplied by a corresponding *a posteriori* probability. Upon a cursory examination of (4) and (5) it would appear to be essentially the same as those previously reported [3]–[5], however, the main difference lies hidden in the computation of the time-varying weighting functions $p(s_{k+1} = s_i \mid Z_{k+1})$.

IV. Calculation of Weighting Coefficients

To recursively compute the *a posteriori* probabilities $p(s_{k+1} = s_i \mid Z_k)$, the following sequence is used: first, the previously stored value $p(s_k = s_\alpha \mid Z_k)$ is updated by a semi-Markov prediction process to $p(s_{k+1} = s_i \mid Z_k)$; then a new measurement is taken and $p(s_{k+1} = s_i \mid Z_{k+1})$ is computed, is used in the adaptive filter, and then is stored to begin the next cycle.

The preceding qualitative procedure can be expressed mathematically by the following set of equations, where use is made of data sequence $Z_{k+1} \triangleq \{Z_k, z_{k+1}\}$. Expanding (3), we have

$$p(s_{k+1} = s_i \mid Z_{k+1}) = \frac{p(z_{k+1} \mid s_{k+1} = s_i, Z_k)p(s_{k+1} = s_i \mid Z_k)}{p(z_{k+1} \mid Z_k)}. \tag{6}$$

The first term, $p(z_{k+1} \mid s_{k+1} = s_i, Z_k)$, can be approximated by a Gaussian density for those cases in which the probability of a transition occurring between any two adjacent time samples is very small. It has been pointed out that the actual density is not Gaussian, but is in fact a weighted sum of Gaussian densities. This is quite true; however, it was determined experimentally from computer simulation that even when the plant randomly switched as often as the duration of several system response times the Gaussian approximation was quite good. In addition, it is pointed out in the Appendix that $p(z_{k+1} \mid s_{k+1} = s_i, Z_k)$ is approximately normally distributed when the probability of a transition occurring between samples is small, and it can be represented by the known Gaussian density function established from the Kalman filtering algorithms conditioned on s_i, i.e., $N\{H\Phi(s_i)\hat{x}_k(s_i), [HM_{k+1}(s_i)H^T + R]\}$.

The second term in the numerator of expression (6), $p(s_{k+1} = s_i \mid Z_k)$, is the predicted value that will be generated by the semi-Markov process. Expanding $p(s_{k+1} = s_i \mid Z_k)$, we find

$$p(s_{k+1} = s_i \mid Z_k) = \sum_{\alpha=1}^{n} p(s_{k+1} = s_i \mid s_k = s_\alpha, Z_k)p(s_k = s_\alpha \mid Z_k). \tag{7}$$

The first term in (7) is conditioned on both $s_k = s_\alpha$ and data sequence Z_k. We have already established that Z_k and $p(s_k = s_\alpha)$ are strongly dependent. In fact, Z_k actually aids in determining $p(s_k = s_\alpha)$, but, since we are given this information, we can express $p(s_{k+1} = s_i \mid s_k = s_\alpha, Z_k)$ by $p(s_{k+1} = s_i \mid s_k = s_\alpha)$. The second term of (7), $p(s_k = s_\alpha \mid Z_k)$, is known from the previous recursive calculation. Combining (6) and (7),

$$p(s_{k+1} = s_i \mid Z_{k+1}) = \frac{\sum_{\alpha=1}^{n} p(z_{k+1} \mid s_{k+1} = s_i, Z_k)p(s_{k+1} = s_i \mid s_k = s_\alpha)p(s_k = s_\alpha \mid Z_k)}{p(z_{k+1} \mid Z_k)}. \tag{8}$$

The only term undefined in (8) is $p(s_{k+1} = s_i \mid s_k = s_\alpha)$, which can be expressed as probability (system is in state i at time t_{k+1} | system is in state α at time t_k), but this is exactly the definition of the "random starting" probability $\theta_{\alpha i}(t_{k+1} - t_k)$ found in the literature dealing with semi-Markov processes [11]. Therefore, (8) can be expressed in its final form by

$$p(s_{k+1} = s_i \mid Z_{k+1}) = \frac{\sum_{\alpha=1}^{n} p(z_{k+1} \mid s_{k+1} = s_i, Z_k)p(s_k = s_\alpha \mid Z_k)\theta_{\alpha i}(t_{k+1} - t_k)}{p(z_{k+1} \mid Z_k)} \tag{9}$$

where the denominator, being common to all terms, acts as a normalization factor and need not actually be computed. It should be pointed out that the form of (9) is not new, only the semi-Markov statistic that is used to make the intermediate predictive step is different. However, this results in a much simpler form of computation and does not require computer storage increasing with increasing time.

In the case of uniform sampling,

$$\theta_{\alpha i}(t_{k+1} - t_k) = \theta_{\alpha i}[(k + 1)T - kT] = \theta_{\alpha i}(T)$$

which depends only upon the sample spacing T. One other item needed in the computation of $p(s_{k+1} = s_i \mid Z_{k+1})$ is the initial probability of being in state s_i at time zero. It was found that the adaptive filter was relatively insensitive to the choice of initial probabilities. Since as data is observed the adaptive filter rapidly learns the true system configuration, one might as well pick equally probable *a priori* estimates with little degradation in expected performance.

V. Example and Simulation

To properly exercise the adaptive state estimator of (4) and (9), a second-order linear system was chosen to randomly switch between three possible (ϕ_i, Γ_i) configurations. In modeling the semi-Markov process it was decided to set all density functions $h_{ij}(\tau)$ equal to 0.07(exp (-0.07τ). This gave a mean switching time $\bar{\tau}_{ij}$ of approximately 14.3 s, which represented about 29 time samples. In other words, the "unreliable" system will be changing state fairly rapidly, making it more difficult to learn and track the true system s_i. Using the semi-Markov design equations for an identically distributed exponential density function, it was found that $\theta_{ij}(T) = 0.007 i \neq j$ and $\theta_{ii} = 0.986$. With this information we are now able to simulate the entire system and filtering algorithms on an IBM 360 digital computer.

In writing a program to simulate the switching system it was assumed that there are three possible plant configurations s_i. A deterministic input of the form $u(k) = 0.8k \exp(-0.033k)$ was also added to the plant input in addition to white noise $w(t)$. The adaptive filter of Fig. 1, consisting of three Kalman filters, was then synthesized with input data $z(t)$ coming from the switching plant.

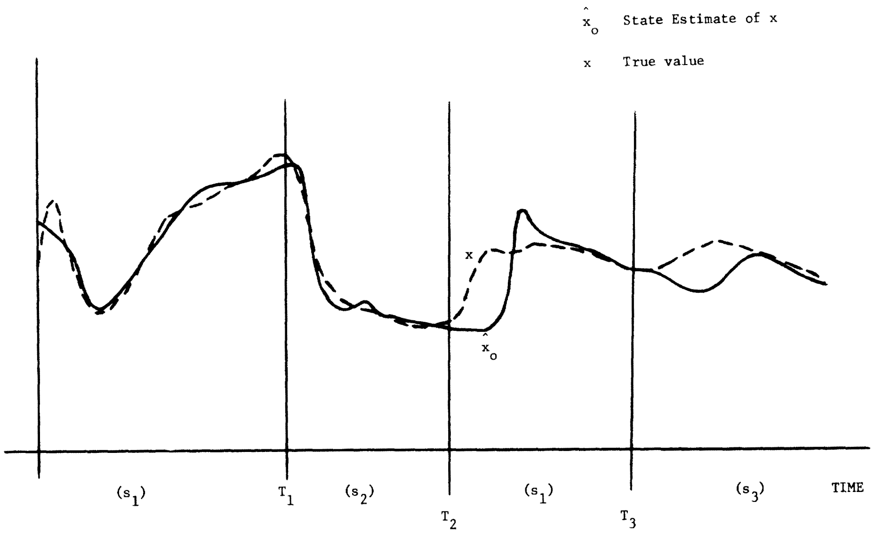

Fig. 2. State estimate $\hat{x}_0$ and true value of x plotted versus time and system transitions T_1, T_2, T_3.

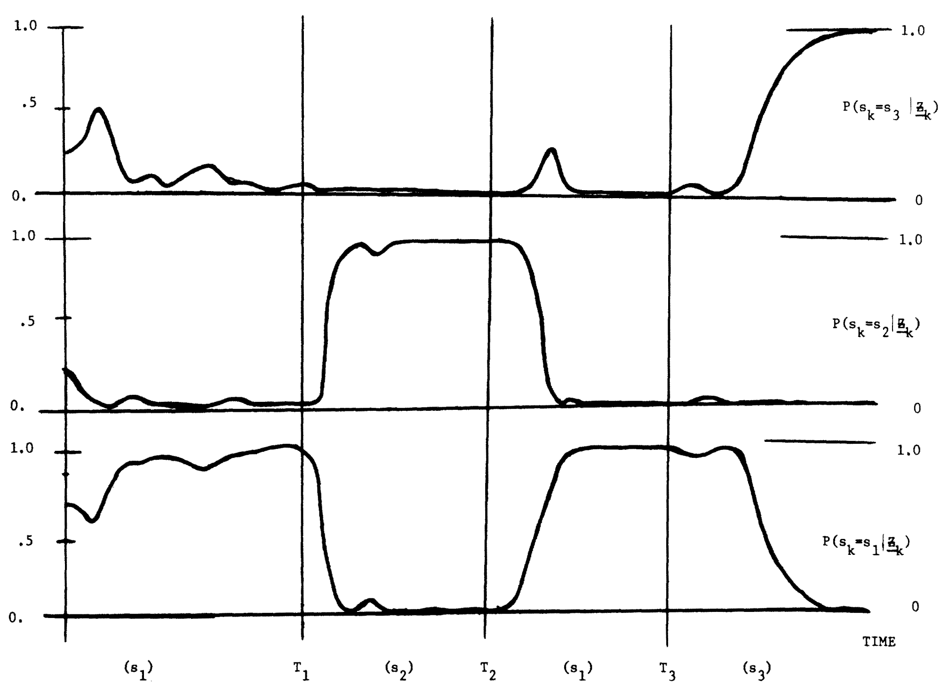

Fig. 3. *Aposteriori* probabilities versus time and system transitions T_1, T_2, T_3.

Results shown in Fig. 2 are for the case of $Q = 1.0$, $R = 4.0$. There are (unknown to the adaptive filter) three plant transitions labeled T_1, T_2, and T_3, respectively, whose time durations are random variables chosen from the exponential density function $h_{ij}(\tau)$. It can be seen from Fig. 2 that the adaptive filter estimate $\hat{x}_0$ (solid line) tracks the true value of state variable x (dotted line) very well. The main source of error in the filter occurs during the learning time (the interval immediately after a plant transition) while the *a posteriori* probabilities make very large transitions.

To better examine what happens during the switching intervals, consider the set of curves shown in Fig. 3, in which the weighting coefficients $p(s_k = s_i \mid Z_k)$ are plotted versus time for $Q = 1.0$, $R = 4.0$. Since the switching sequence is s_1 to s_2 to s_1 to s_3, the probability curve $p(s_k = s_i \mid Z_k)$ should be close to unity during the period of time when the system is truly in state s_i and should approach zero at all other times. Fig. 3 shows that indeed this is the case with probabilities between 0.90 and 0.999 once the true system is learned.

VI. Conclusion

A nonlinear state estimator based upon Kalman–Bucy theory and semi-Markov statistics has been developed to help solve the switching plant problem where the plant configuration randomly switches between a finite set of possible configurations at unknown times. By utilizing the semi-Markov process to model the switching plant and incorporating the process into the design of an adaptive filter, the problem of growing computer storage was completely eliminated. The design procedure was illustrated by the analysis of an example chosen to simulate a rapidly switching plant in the presence of a low signal-to-noise ratio (large measurement error). Results were analyzed, and the performance, with the exception of a short period of learning, was shown to approach that of an unrealizable optimal filter whose parameters automatically change with each change in the plant. It should be emphasized that the adaptive state estimator is suboptimal due to the restrictive assumptions placed upon its design. The restrictions are as follows: 1) that the filter of (5) can be realized by a Kalman filter, and 2) that the unreliable system switch slowly enough such that the first term of design equation (9) can be represented by a Gaussian density function. In practice this is not nearly as restrictive as it sounds since many unreliable systems change configuration fairly slowly with respect to their normal system response time.

Appendix

Approximation of $p(z_{k+1} \mid s_{k+1} = s_i, Z_k)$

There are several possible approaches that can be taken to obtain an accurate approximation for the term $p(z_{k+1} \mid s_{k+1} = s_i, Z_k)$ appearing in the *a posteriori* weighting equation (9). The approach chosen here will employ Bayes' rule and assume that the system parameters change slowly with respect to the sampling rate. An assumption generally valid for many practical purposes, as can be seen from the example given here, is $\phi_{ii} = 0.986$ and $\phi_{ij} = 0.007$. Thus, ϕ_{ij} can be approximated by the Kronecker delta function δ_{ij}.

By Bayes' rule, we have

$$p(z_{k+1} \mid s_{k+1} = s_i, Z_k) = \sum_{\alpha=1}^{n} p(z_{k+1} \mid s_k = s_\alpha, s_{k+1} = s_i, Z_k) p(s_k = s_\alpha \mid s_{k+1} = s_i, Z_k). \tag{A-1}$$

Now, the first term is equal to the normal distribution

$$N[H_{k+1}\phi(s_\alpha)\hat{x}_k(s_\alpha), (H_{k+1}M_{k+1}(s_\alpha)H_{k+1}^T + R)]$$

while the second term can be approximated by the Kronecker delta $\delta_{i\alpha}$. Equation (A-1) can then be rewritten

$$p(z_{k+1} \mid s_{k+1} = s_i, Z_k) \simeq \sum_{\alpha=1}^{n} p(z_{k+1} \mid s_k = s_\alpha, s_{k+1} = s_i, Z_k)\delta_{i\alpha} = p(z_{k+1} \mid s_k = s_i, s_{k+1} = s_i, Z_k)$$

which represents the normal distribution

$$N[H_{k+1}\phi(s_i)\hat{x}_k(s_i), (H_{k+1}M_{k+1}(s_i)H_{k+1}^T + R)].$$

Acknowledgment

The authors wish to thank the reviewers for their valuable comments.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, Ser. D, vol. 82, pp. 35–45, Mar. 1960.
[2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Eng.*, Ser. D, vol. 83, pp. 95–108, Mar. 1961.
[3] D. T. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 434–439, Oct. 1965.
[4] C. G. Hilborn, Jr., and D. G. Lainiotis, "Optimal estimation in presence of unknown parameters," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-5, pp. 38–43, Jan. 1969.
[5] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 10–16, Feb. 1970.
[6] C. G. Hilborn, Jr., and D. G. Lainiotis, "Learning systems for minimum risk adaptive pattern classification and optimal adaptive estimation," Dep. Elec. Eng., Univ. Texas, Austin, CSRG Tech. Rep. 67-9, DDC Document AD-663702, Nov. 1967.
[7] D. G. Lainiotis, "Supervised learning sequential structure and parameter adaptive pattern recognition: discrete data case," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 106–110, Jan. 1971.
[8] ——, "Sequential structure and parameter-adaptive pattern recognition—part I: supervised learning," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 548–556, Sept. 1970.
[9] H. B. Parekh, "A study and application of estimation under uncertainty," Ph.D. dissertation, Southern Methodist Univ., Dallas, Tex., 1970.
[10] H. B. Parekh and J. L. Melsa, "Optimal and suboptimal estimation with stochastic parameters," in *Proc. Mervin J. Kelly Conf.*, Oct. 1970.
[11] R. A. Howard, "System analysis of semi-Markov processes," *IEEE Trans. Military Electronics*, vol. MIL-8, pp. 114–124, Apr. 1964.

Absolutely Expedient Learning Algorithms for Stochastic Automata

S. Lakshmivarahan and M. A. L. Thathachar

***Abstract*—A general nonlinear reinforcement scheme of the reward-penalty type for a multistate stochastic automaton acting in stationary random media is described. A general but simple condition of symmetry of the nonlinear functions figuring in the reinforcement scheme is shown to be necessary and sufficient for absolute expediency (monotonic decrease of the expectation of the average penalty in all stationary random media). Various schemes are simulated and the results are compared. An adaptive updating of the parameters in the scheme to obtain faster convergence is proposed.**

I. Introduction

The stochastic automaton is a discrete finite-state device having $m(\geq 2)$ states W_i, $i = 1,2,\cdots,m$. At any instant of time the automaton, being in state $W(n) = W_i$, is governed by the state probability distribution $\pi_i(n)$, where n is the discrete time variable which takes on values $0,1,2,\cdots$ such that at all times

$$\sum_{i=1}^{m} \pi_i(n) = 1. \tag{1}$$

When in state W_i the automaton gives an output or action Y_i. The response $S_i(n)$ of the medium or environment to this action Y_i appearing at the input to the automaton can take on two values: $S_i(n) = 1$, known as the penalty input, with probability p_i; and $S_i(n) = 0$, known as the nonpenalty input, with probability $q_i = 1 - p_i$. For simplicity, the p_i, $i = 1,\cdots,m$, are called the penalty probabilities of the medium. It is assumed that there exist unique p_l and p_k such that

$$p_l = \min_j \{p_j\} \tag{2}$$

$$p_k = \max_j \{p_j\} \tag{3}$$

and that the p_j do not vary with time, that is, the medium is stationary random. Except for the fact that there exist unique p_l and p_k satisfying (2) and (3), the actual values of the p_j and the values of l and k are generally unknown. The average penalty the automaton receives from the medium is

$$M(n) = \sum_{i=1}^{m} \pi_i(n)p_i = \boldsymbol{\pi}^T(n)\boldsymbol{p} \tag{4}$$

Manuscript received January 20, 1972; revised October 10, 1972.
The authors are with the Department of Electrical Engineering, Indian Institute of Science, Bangalore, India.